CHAPTER 5

Electric and Magnetic Sensors and Actuators

Torpedoes, sharks, eels, pigeons, magnetic bacteria, and the platypus

Electric and magnetic fields are too important and too common to be neglected by nature in its grand design. Many animals and organisms have found ways to take advantage of these fundamental forces for sensing and actuation. The electric field in particular is used for both sensing and actuation. Almost all rays and sharks can sense electric fields produced by prey, as can some catfish, eels, and the platypus. Electric fields are sensed through use of special gelatinous pores that form electroreceptors called ampulae of Lorenzini. Sensing can be passive or active. Sharks and rays use passive sensing; prey is located by sensing weak electric fields produced by the muscles and nerves in the prey. Some animals, such as the electric fish, can generate electric fields for the purpose of active electrolocation of prey. The same basic sensory system is used by young sharks for protection by freezing in place when electrolocation fields are detected. But perhaps the best known example of electrolocation is the platypus, which uses electroreceptors in its bill to hunt by night. Actuation is just as common and is used primarily to stun prey, and also for protection. The torpedo or electric ray (genus *Torpedinidae*) is one of some 70 species of rays that can produce voltages and apply them in a manner similar to a battery. The voltage is produced in a pair of electric organs made of plates connected to a nervous system that controls them. In rays, these biological batteries are connected in parallel to produce low-voltage, high-current sources. The range is between 8 V to more than 200 V with currents that can reach a few amperes. Another example is the electric eel (*Electrophorus electricus*). Since it lives in freshwater, which is less conductive than seawater, it has its plates in series to produce higher voltages (up to 600 V at perhaps 1 A, in short pulses).

Magnetic fields are equally important in sensing. It is now well established that many birds can sense the terrestrial magnetic field and use this ability for navigation. Pigeons are known to have a biocompass made of magnetite particles in the upper tissue of their beaks and use that for magnetolocation. Traces of magnetite can be found even in the human brain, suggesting that perhaps we also had this ability in the distant past. Even bacteria have found a use for magnetite—to help them move along the lines of the terrestrial magnetic field. The bacterium *Magnetospirilllum magneticum* uses magnetite particles to orient itself along lines of the magnetic field to allow them to reach environments rich in oxygen.

5.1 INTRODUCTION

The class of electric and magnetic sensors and actuators is the broadest by far of all other classes, both in numbers and types and the variety within each type. Perhaps this should come as no surprise since in a majority of cases a sensor exploits the electrical properties of materials and, with few exceptions, the requisite output is electrical. In fact, we could argue that even sensors that were not placed in this category belong here as well. Thermocouples exploit electrical effects in conductors and semiconductors—an electrical phenomenon. Optical sensors are either based on wave propagation, which is an electromagnetic phenomenon, or on quanta, which are measured through electronic interaction with the atomic structure of the sensor. It would take little to argue that this is an electric phenomenon. In terms of actuation, most actuators are either electrical or, more commonly, magnetic. This is particularly true of actuators that need to provide considerable power. However, for the sake of simplicity and to follow the basic idea of limiting the number of principles involved in each class of sensors, we will limit ourselves here to the following types of sensors and actuators:

1. Sensors and actuators based on electric and electrostatic principles, including capacitive sensors (proximity, distance, level, material properties, humidity, and other quantities such as force, acceleration, and pressure) and related electric field sensors and actuators. This class of sensors includes microelectromechanical (MEMs) devices, but these will be discussed separately in **Chapter 10**.
2. Sensors based on direct measurement of resistance. There are many sensors that belong in this category, including both AC and DC sensing of current and voltage, position and level sensing, and many others.
3. Magnetic sensors and actuators based on the static and low-frequency time-dependent magnetic field. The variety here is large. It includes motors and valves for actuation, magnetic field sensors (hall element sensors, inductive sensors for position, displacement, proximity, etc.), and a variety of others, including magnetostrictive and magnetoresistive sensors and actuators.

A fourth group of electromagnetic sensors, based on radiation effects of the electromagnetic field will be discussed in **Chapter 9**.

One can classify the sensors and actuators discussed here as electric, magnetic, and resistive. Often we will simply call them electromagnetic devices, a term that encompasses all of them (as well as those discussed in **Chapter 9**).

All electromagnetic sensors and actuators are based on electromagnetic fields and their interaction with physical media. It turns out that the electric and magnetic fields in media are influenced by or influence a large number of properties. For this reason, electromagnetic sensors for almost any imaginable quantity or effect either exist or can be designed.

Before continuing it is perhaps useful to bear in mind the following definitions:

An electric field is the force per unit charge that exists in the presence of charges or charged bodies. An electric field may be static when charges do not move or move at constant velocity or may be time dependent if charges accelerate and/or decelerate.

Moving charges in conducting media or in space cause currents and currents produce magnetic fields. Magnetic fields are either static, when currents are constant (DC), or time dependent, when currents vary in time.

When currents vary in time, both an electric and a related magnetic field are established. This is called an electromagnetic field. Strictly speaking, an electromagnetic field implies that both an electric and a magnetic field exist. However, it is not entirely wrong to call all electric and magnetic fields by that name since, for example, an electrostatic field may be viewed as a time-independent electromagnetic field with zero magnetic field. Although the properties of the various fields are different, they are all described by Maxwell's equations.

Electromagnetic actuators are based on one of the two basic forces: electric force (best understood as the attraction between opposite polarity charges or repulsion between like polarity charges) and magnetic force. The latter is the attraction of conductors carrying currents in the same direction or repulsion of current carrying conductors with currents in opposite directions. Perhaps the best-known manifestation of this force is between two permanent magnets.

5.2 | UNITS

The basic SI electric unit is the ampere (A), as was discussed in **Section 1.6**. But the subject of electricity and magnetism includes a fairly large number of derived units based on the laws and relations in electromagnetics. The current itself is sometimes specified as a density either as current per unit area (ampere/meter^2 [A/m^2]) or in some cases as a current per unit length (ampere/meter [A/m]). These are current densities. Aside from the ampere as a unit of electric current, probably the most common unit is the volt, defined as energy per unit charge. Charge itself has units of coulombs (C) or ampere-seconds (A·s), but charge is often encountered as a density: coulomb/meter (C/m), coulomb/square meter (C/m^2), or coulomb/cubic meter (C/m^3). The derived unit for voltage is indicated as the volt (V) or as joules/coulomb (J/C), newton-meter/coulomb (N·m/C), or, as was shown in **Section 1.6**, the purely SI form is kg·m^2/A·s^3. The product of voltage and current is power, measured in watts (W), and although power can be described in other units (such as force $\times$ velocity, i.e., N·m/s or J/s), in electrical engineering it is rare to use any other unit except the watt or the ampere-volt (A·V). Work and energy are normally measured in joules (J), but in measuring and describing electrical energy it is customary to use the basic unit of watt $\times$ time, such as the watt-hour (W·h) or one of its multipliers such as kW·h. Energy density (per unit volume) is used to signify energy storage capacity in joules/cubic meter (J/m^3).

The ratio between voltage and current (V/A) is **resistance** (Ohm's law), so the unit of resistance is the ohm (Ω). The **conductivity** of a medium is one of the three fundamental electric properties of any medium, the other two being the **electric permittivity** and **magnetic permeability**. Conductivity is easier to understand by first defining its reciprocal, called **resistivity**, which is the measure of resistance of the material, given in ohm-meters (Ω·m). The conductivity, which is a measure of how well the medium conducts current, has units of 1/(Ω·m) and is the reciprocal of resistivity. The unit 1/Ω is known as the siemens (S). Hence the unit of conductivity is the siemens/meter (S/m). The ratio between charge and voltage is called **capacitance** and is indicated as the farad (F) (see **Example 1.4**). In addition, one encounters the **electric field intensity** in units of volts/meter (V/m) or newtons/coulomb (N/C). Permittivity is measured in farads/meter (F/m), a unit that can be easily derived from Coulomb's law (to be discussed later in this chapter). One can also define an **electric flux density** as the product of the electric field intensity and permittivity with units of coulombs/square meter (C/m^2) and, by integration of the **electric flux** density over an area, an electric flux with units of coulombs (C).

The magnetic field is typically given in terms of the **magnetic flux density** (sometimes called induction) or the **magnetic field intensity**. The derived unit for the magnetic field intensity is amperes/meter (A/m), derived from Ampere's law. The flux density is measured in units of tesla (T). The tesla is in fact force per unit length per unit current (N/A/m or kg/A·s^2). A commonly used nonmetric unit of magnetic flux density is the gauss (g), (1 T = 10,000 g). Integration of the magnetic flux density over an area produces a **magnetic flux** with units of tesla square meter (T·m^2), designated as the weber (Wb). Therefore the magnetic flux density can also be indicated as weber/square meter (Wb/m^2), explicitly showing the fact that it is a density. The ratio of magnetic flux and current is **inductance**, whose unit is webers/ampere (Wb/A) or tesla square meter/ampere (T·m^2/A), also known as the henry (H).

The ratio between magnetic flux density and magnetic field intensity is the magnetic permeability of the medium. Its units are clearly T·m/A. However, the quantity T·m^2/A is designated as the henry and hence permeability has units of henrys/meter (H/m). Another magnetic quantity of limited use is the **magnetic reluctance** or **reluctivity**, which may be best understood as a kind of "magnetic conductivity" with units of 1/henry (1/H). In addition to the above, one should also recall that the frequency of a signal is indicated in cycles/second (cycles/s) or hertz (Hz). Related to frequency is the angular frequency (also called angular velocity) in radians/second (rad/s).

There are other quantities that are sometimes used, such as the phase of a signal (degrees or radians), attenuation and phase constants, power density, and more, but these are best defined in the context of their use (some of these quantities will be introduced in **Chapter 9**). Some of the quantities described above (current density, electric and magnetic field intensities, and electric and magnetic flux densities) are vectors that have both a magnitude and direction, the rest are scalars and only posses a magnitude. Power can be represented as a scalar or a vector, but we will view it here strictly as a scalar. **Table 5.1** summarizes the quantities and units of many of the electric and magnetic quantities.

5.3 | THE ELECTRIC FIELD: CAPACITIVE SENSORS AND ACTUATORS

Electric field sensors and actuators are those that operate on the physical principles defining the electric field and its effects. The primary type of device is capacitive. The discussion here is in terms of capacitances because this affords a simple circuit approach, but it can equally well be done in terms of the electric field intensity. There are some sensors, such as charge sensors, that are better explained in terms of the electric field, but on the whole, discussion of capacitance and its use in sensing and actuation covers most aspects necessary for a thorough understanding of the principles involved without the need to study the intricacies of the electric field behavior.

All capacitive sensors are based on the change in capacitance due to the stimulus either directly or indirectly. First, it should be noted that capacitance, by definition, is the ratio between charge and voltage on a device:

$$C = \frac{Q}{V} \quad [\mathrm{C/V}]. \tag{5.1}$$

Capacitance is measured in coulombs/volt (C/V). This unit is called the farad (F). Because voltage is only properly defined as the difference in potentials between two

TABLE 5.1 ■ Electric and magnetic quantities and their units

	Unit	Notes	Symbol
Current	ampere (A)	(SI unit)	A
Current density	ampere/meter^2 (A/m^2), ampere/meter (A/m) (see text)	vector	J
Voltage	volt (V)	also emf	V
Charge	coulomb (C), (A·s)		Q, q
Charge density	coulombs/meter (C/m), coulomb/square meter (C/m^2), coulomb/cubic meter (C/m^3)		ρ_1, ρ_s, ρ_v
Permittivity	farad/meter (F/m)		ε
Electric field intensity	volt/meter (V/m) or newton/coulomb (N/C)	vector	E
Electric flux density	coulomb/square meter (C/m^2)	vector	D
Electric flux	coulomb (C)		Φ or Φ_e
Power	watt (W), (A·V)		P
Energy	watt-hour (W·h), joule (J)		W
Energy density	joule/cubic meter [(J/m^3)		w
Resistance	ohm (Ω)		R
Resistivity	ohm-meter (Ω·m)		ρ
Conductivity	siemens/meter (S/m) or 1/ohm-meter (1/Ω·m)	$\sigma = 1/\rho$	σ
Capacitance	farad (F)		C
Magnetic flux density	tesla (T), sometimes gauss (g), weber/square meter (Wb/m^2)	vector	B
Magnetic field intensity	ampere/meter (A/m)	vector	H
Magnetic flux	tesla-square meter (T·m^2) or weber (Wb)		Φ or Φ_m
Inductance	weber/ampere (Wb/A) or henry (H)		L
Magnetic permeability	henry/meter (H/m)		μ
Reluctance	1/henry (1/H)		R
Frequency	cycles/second (cycle/s) or hertz (Hz)		f
Angular frequency	radians/second (rad/s)	$\omega = 2\pi f$	ω

points, the capacitance is only defined for two conducting bodies across which the potential difference is connected. This is shown schematically in **Figure 5.1**. Body *B* is charged by the battery to a positive charge *Q* and body *A* to an equal but negative charge $-Q$. Any two conducting bodies, regardless of size and distance between them, have a capacitance. The capacitance of a single conducting body can also be defined in terms of the charge of that body and potential difference with respect to infinity as a special case of a two-body system. When a potential difference is connected across them, the bodies are charged and the relation in **Equation (5.1)** is satisfied. Nevertheless, capacitance is independent of voltage or charge—it only depends on physical dimensions and materials properties. To understand the principles, we start with the

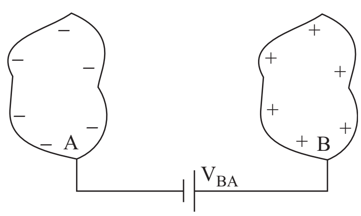


FIGURE 5.1 ■ The concept and definition of capacitance.

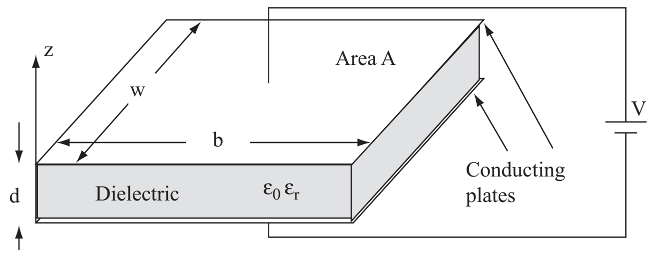


FIGURE 5.2 ■ The parallel plate capacitor connected to a DC source.

capacitance between two parallel plates, shown in **Figure 5.2**. Assuming first that the distance d between the two plates is small, the capacitance of the device is

$$C = \frac{\varepsilon_0 \varepsilon_r S}{d} \quad [\mathrm{F}], \tag{5.2}$$

where ε_0 is the permittivity of a vacuum, ε_r is the relative permittivity (dielectric constant) of the medium between the plates, S is the area of the plates, and d is the distance between the plates. ε_0 is a constant equal to 8.854×10^{-12} F/m, whereas ε_r is the ratio between the permittivity of the medium and that of free space (ε_0) and hence is dimensionless. Permittivity is measurable and available as part of the electrical properties of materials. Although not strictly necessary for the definition, it is usually assumed that the material between the plates of the capacitor is a nonconducting medium—a dielectric. The relative permittivities of some common dielectrics are listed in **Table 5.2**.

Any of the quantities in **Equation (5.2)** affect the capacitance and changes in these can be sensed. This allows for a wide range of stimuli, including displacement and anything that can cause displacement (pressure, force), proximity, permittivity (e.g., in moisture sensors), and a myriad of others. However, **Equation (5.2)** describes a very specific device and was obtained by assuming that the electric field intensity between the two plates does not leak (fringes) outside the space between the plates. This was done to obtain a simple expression. In the more general case, when d is not small, or if the plates are arranged in a different configuration (see **Figure 5.3**), we cannot calculate the capacitance directly, but we can still write the following:

$$C = \alpha[\varepsilon_0, \varepsilon_r, S, 1/d]. \tag{5.3}$$

TABLE 5.2 ■ Relative permittivities of various materials

Material	ε_r	Material	ε_r	Material	ε_r
Quartz	3.8–5	Paper	3.0	Silica	3.8
Gallium arsenide	13	Bakelite	5.0	Quartz	3.8
Nylon	3.1	Glass	6.0 (4–7)	Snow	3.8
Paraffin	3.2	Mica	6.0	Soil (dry)	2.8
Perspex	2.6	Water (distilled)	81	Wood (dry)	1.5–4
Polystyrene foam	1.05	Polyethylene	2.2	Silicon	11.8
Teflon	2.0	Polyvinyl chloride	6.1	Ethyl alcohol	25
Barium strontium titanate	10,000.0	Germanium	16	Amber	2.7
Air	1.0006	Glycerin	50	Plexiglas	3.4
Rubber	3.0	Nylon	3.5	Aluminum oxide	8.8

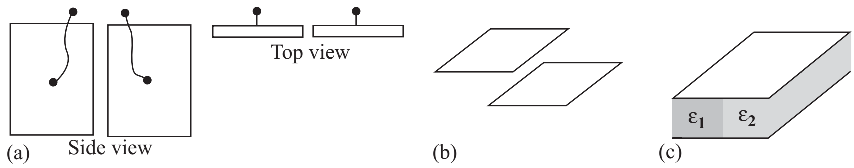


FIGURE 5.3 ■ Various configurations of parallel plate capacitors.

That is, capacitance is proportional to permittivity and the area of the conductors (plates) and inversely proportional to the distance between conductors. The parallel plate capacitor is only one possible device. As long as two conductors are involved, there will be a definable capacitance between them. **Figure 5.4** shows some other useful capacitive arrangements often exploited for sensing. Many capacitive sensors will be encountered in **Chapters 6–10**, but we will discuss here the sensing of position, proximity, displacement, and fluid level, as well as a number of capacitive actuators.

5.3.1 Capacitive Position, Proximity, and Displacement Sensors

Returning to **Equation (5.2)**, position and displacement can be used to change the capacitance of a device in three fundamental ways:

1. By allowing one conductor in a two-conductor capacitor (usually a plate) to move relative to the other. A number of configurations are shown in **Figure 5.5**. In **Figure 5.5a**, the sensor is made of a single plate while the second plate is a conductor relative to which the distance (proximity) is sensed. While this is a valid method, it is not a sensor one can obtain ready-made, but rather one has to build it, and this implies that proximity can only be sensed with respect to a conducting surface. A schematic position sensor of this type is shown in **Figure 5.6**. One plate

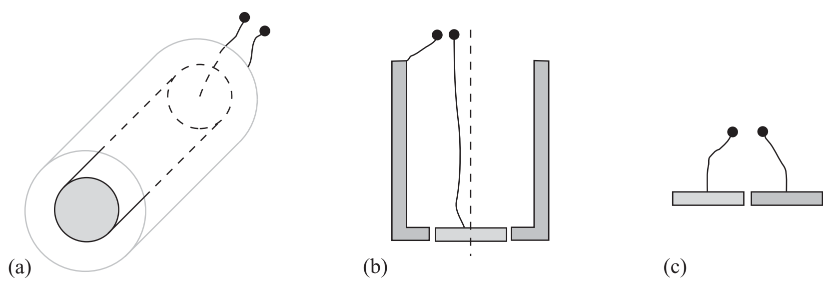


FIGURE 5.4 ■ Various types of capacitive sensor arrangements.

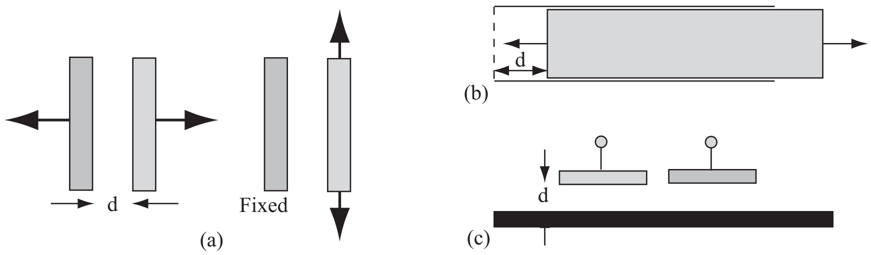


FIGURE 5.5 ■ Arrangements of capacitive sensors for position, proximity, and displacement sensing. (a) One plate is usually fixed and the measurand is the change in distance *d* or surface area *S* of the capacitor. (b) Change in permittivity. (c) Change in distance.

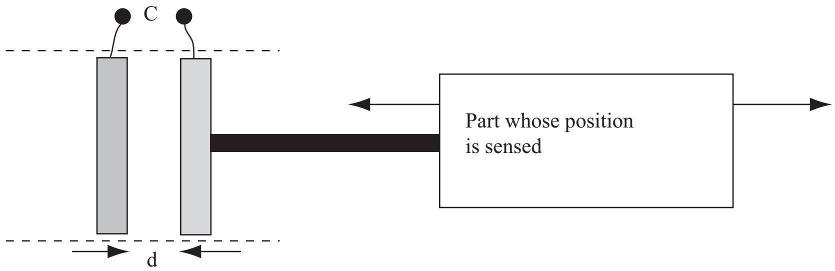


FIGURE 5.6 ■ A schematic capacitive position sensor.

is fixed while the other is pushed by the moving device. The position of the moving device causes a change in position of one of the plates, and this changes the capacitance. The capacitance is inversely proportional to distance, and as long as the distances sensed are small, the output is linear.

2. Alternatively, the plates remain fixed but the dielectric moves in or out, as in **Figure 5.5b**. This is a practical situation for some applications. For example, the dielectric may be connected to a float that then senses the fluid level, or it may be pushed by a device to sense the end of travel or position. The advantage of this device is that it is quite linear, and the range of motion is rather large since it can approach the width of the capacitor.
3. Another configuration is obtained by keeping the plates fixed, as in **Figure 5.5c**, and sensing the distance to a surface. This is a practical arrangement since the sensor is self-contained and requires no external electrical connections or physical arrangements to sense distance or position. However, the relation between capacitance and distance is nonlinear and distance is limited because the electric field does not extend very far.

EXAMPLE 5.1 Small capacitive displacement sensor

A small sensor capable of accurate displacement sensing can be built from two small plates, as in **Figure 5.5a**. The plates can move either toward each other or slide sideways. The sensors discussed here are as follows:

a. The two plates are 4 mm × 4 mm and move toward or away from each other with a minimum displacement of 0.1 mm and a maximum displacement of 1 mm (**Figure 5.7a**).

b. The two plates are 4 mm × 4 mm and are separated a fixed distance of 0.1 mm. They slide sideways with a displacement range of 0–2 mm (**Figure 5.7b**).

Solution:

a. The capacitance is calculated using a variable distance d (0.1 mm $< d <$ 1 mm) in **Equation (5.2)** based on **Figure 5.7a**. This is shown in the first row of the following table. Capacitance is in picofarads (pF). Because the plates are small, the capacitance is also small and the effects of the edges are likely to introduce errors in calculation of the capacitance using the parallel plate capacitor formula. To see what this effect is, the capacitance is also calculated numerically using a method called the method of moments, which allows exact computation of the capacitance and does not require the approximate capacitance formula in **Equation (5.2)**. The second row shows

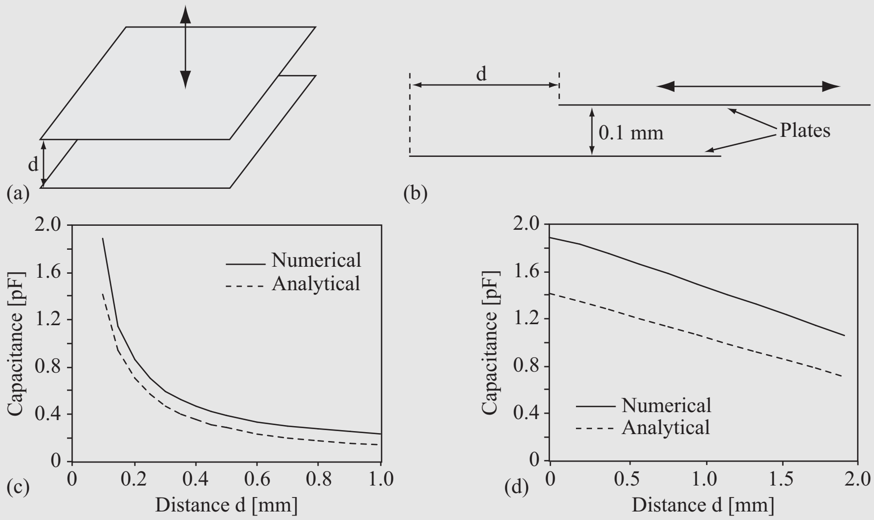


FIGURE 5.7 ■ Capacitive displacement sensors. (a) Lower plate fixed, upper plate moves up and down to indicate position. (b) Lower plate fixed, upper plate moves sideways to indicate position. (c) Capacitance as a function of displacement (analytical and numerical) for the sensor in (a). (d) Capacitance as a function of displacement for the sensor in (b).

the results obtained using the method of moments (the same can be done experimentally using a capacitance meter).

d (mm)	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.6	0.7	0.8	0.9	1.0
C (pF)	1.89	1.15	0.862	0.7	0.595	0.520	0.465	0.422	0.388	0.337	0.3	0.273	0.251	0.234
C (pF)	1.42	0.944	.708	0.567	0.472	0.405	0.354	0.315	0.283	0.236	0.202	0.177	0.157	0.142

b. The capacitance is calculated using a variable horizontal offset of the upper plate x, $(0.0 \text{ mm} < x < 2 \text{ mm})$ in **Equation (5.2)**. This is shown in the first row of the following table. Capacitance is in picofarads. The second row shows the results obtained using the method of moments.

x (mm)	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
C (pF)	1.89	1.83	1.75	1.67	1.58	1.49	1.41	1.32	1.23	1.15	1.06
C (pF)	1.42	1.35	1.27	1.2	1.13	1.06	0.992	0.921	0.850	0.779	0.708

Clearly, the analytical results using **Equation (5.2)** and the numerical (or experimental) results seem to be very different, as can be expected from the small plates (the larger the plates and the smaller the distance between them, the closer the analytical and numerical results become). However, a plot of the two results reveals additional information. The plots of the data in the two tables are shown in **Figures 5.7c** and **5.7d**. The first thing to note is that the two sets of results in each figure are essentially the same in shape, but shifted with respect to each other. The second piece of information is that the lateral moving sensor is more linear but less sensitive (the change in capacitance is smaller). Both can be used successfully once properly calibrated, but it should be noted again that for the sensors shown here, the numerical (or experimental) data are appropriate to use whereas the analytical calculation gives the general behavior but not useful, accurate data.

In most proximity sensors, the method in **Figure 5.5c** is the most practical, but the construction is somewhat different. One typical type of sensor is made as follows: a hollow cylindrical conductor forms one plate of the sensor, as in **Figure 5.8**. The second plate of the sensor is a disk at the lower opening of the cylinder. The whole structure may be enclosed with an outer conducting shield or may be encased in a dielectric enclosure. The capacitance of the device is C_0 based on dimensions, materials, and structure. When any material is present in the proximity of the lower disk, it changes the effective permittivity seen by the sensor and its capacitance increases to indicate the distance between the sensor and the surface. The advantage of this sensor is that it can sense distances to conducting or nonconducting bodies of any shape, but the output is not linear. Rather, the smaller the sensed distance d, the larger the sensitivity of the sensor. The dimensions of the sensor have a large influence on its span and sensitivity. Large-diameter sensors will have a larger span but relatively low sensitivity, whereas smaller diameter sensors will have a smaller span with greater sensitivity. **Figure 5.9** shows some capacitive proximity sensors (of different physical sizes and sensing distances) that may be used to sense conducting surfaces and/or switch on at a preset distance.

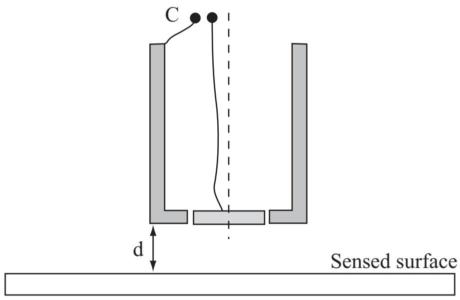


FIGURE 5.8 ■ A practical proximity sensor arrangement.

FIGURE 5.9 ■ Capacitive proximity sensors.

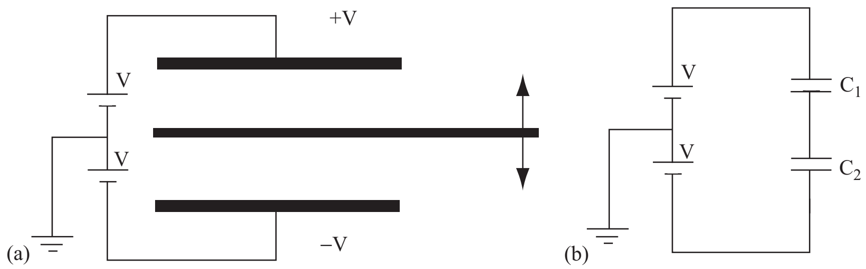


FIGURE 5.10 ■ Position sensor arrangement with improved linearity. (a) Sensor. (b) Equivalent circuit.

Capacitive position and proximity sensors may be made in other ways. One example is the sensor shown in **Figure 5.10**. The sensor has two fixed plates and one moving plate. When the plate is midway, its potential is zero with respect to ground since $C_1 = C_2$. As the plate moves up, its potential becomes positive (C_1 increases, C_2 decrease). When it moves down it is negative (C_2 increases, C_1 decreases). A sensor of this type tends to be more linear than the previous sensors, but the distance between the fixed plates must be small and consequently the motion must also be small or the capacitances will be very small and difficult to measure. Other capacitive sensors sense rotary motion by rotating one plate with respect to the other. Still others are cylindrical or made in any convenient shape such as a comb shape (see **Figure 5.11**).

5.3.2 Capacitive Fluid Level Sensors

Fluid level may be sensed by any of the position or proximity sensors discussed in the previous paragraph by sensing the position of the fluid surface either directly or through a float that then can change the capacitance of a linear or rotary capacitor. One of the simplest, direct methods for fluid level sensing is to allow the fluid to fill the space between the two conducting surfaces that make up the capacitor. For example, the capacitance of the parallel plate capacitor is linearly proportional to the permittivity between the two plates. Therefore the larger the amount of water between the plates, the larger the capacitance and therefore the capacitance is a measure of fluid level between the plates. **Figure 5.12** shows a parallel plate capacitor used as a fluid level sensor. The part of the plates under the surface of the water has a capacitance C_f:

$$C_f = \frac{\varepsilon_f hw}{d} \quad [\mathrm{F}], \tag{5.4}$$

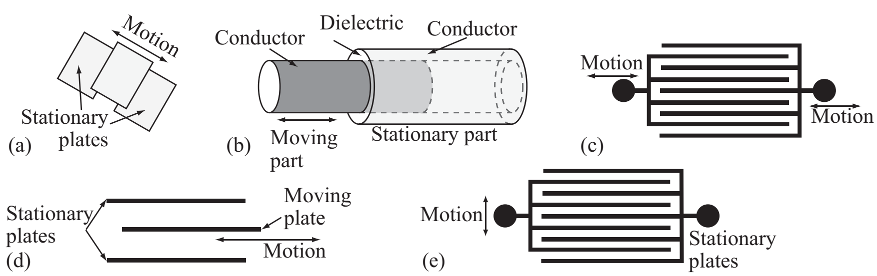


FIGURE 5.11 ■ Various arrangements for linear displacement capacitive sensors.

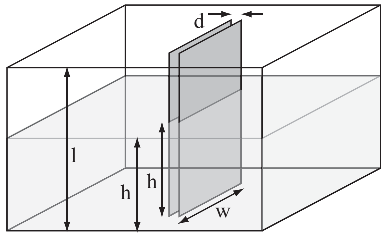


FIGURE 5.12 ■ The principle of a capacitive fluid level sensor. The fluid must be nonconducting.

where ε_f is the permittivity of the fluid, h is the height of the fluid, w is the width of the plates, and d is the distance between them. The part of the capacitor above the fluid has a capacitance C_0:

$$C_0 = \frac{\varepsilon_0(l-h)w}{d} \quad [\mathrm{F}], \tag{5.5}$$

where l is the total height of the capacitor. The total capacitance of the sensor is the sum of the two:

$$C = C_f + C_0 = \frac{\varepsilon_f hw}{d} + \frac{\varepsilon_0(l-h)w}{d} = h\left[\frac{(\varepsilon_f - \varepsilon_0)w}{d}\right] - \frac{\varepsilon_0 lw}{d} \quad [\mathrm{F}]. \tag{5.6}$$

Clearly this relation is linear and varies from the minimum capacitance $C_{\min} = \varepsilon_0 lw/d$ (for $h = 0$) to a maximum capacitance $C_{\max} = \varepsilon_f lw/d$ (for $h = l$). The sensitivity of the sensor may be calculated as dC/dh and is clearly linear.

Although a parallel plate capacitor can be used for this purpose, the expressions above are approximate (they neglect the effect of the edges by assuming the field is not affected by the finite size of the plates). In reality there will be a slight nonlinearity due to these effects and this nonlinearity depends on the distance between the plates. Also, it should be noted that the method is only practical with nonconducting fluids (oils, fuels, freshwater). For slightly conducting fluids, the plates must be coated with an insulating medium.

A more common implementation of this simple, rugged sensor is shown in **Example 5.2**.

EXAMPLE 5.2 Capacitive fuel gauge

A fuel tank gauge is made as shown in **Figure 5.13a**. A long capacitor is made from two coaxial tubes immersed in the fuel so that the fuel fills the space between them up to the fluid level. The empty cylinders (empty fuel tank) establish a capacitance C_0. The capacitance of a coaxial capacitor of length d, inner radius a, and outer radius b is given by

$$C_0 = \frac{2\pi\varepsilon_0 d}{\ln(b/a)} \quad [\mathrm{F}].$$

If the fluid fills the capacitor to a height h, the capacitance of the device is

$$C_f = \frac{2\pi\varepsilon_0}{\ln(b/a)}(h\varepsilon_r + d - h) = \frac{2\pi\varepsilon_0}{\ln(b/a)}(\varepsilon_r - 1)h + \frac{2\pi\varepsilon_0}{\ln(b/a)}d \quad [\mathrm{F}],$$

where ε_r is the relative permittivity of the fuel. Clearly then the capacitance is linear with respect to h from $h = 0$ to $h = d$.

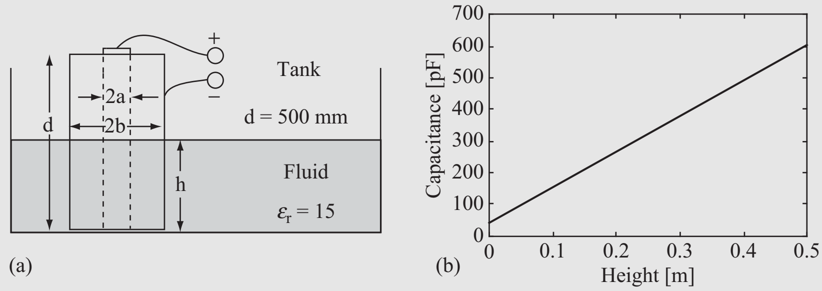


FIGURE 5.13 ■ A fluid level sensor (or fuel gauge) with improved transfer function. (a) The sensor. (b) The transfer function for the values given. For the best performance, the distance between the inner and outer conductor should be small to reduce edge effects and hence possible nonlinearities at very low and very high levels due to fringing at the edges.

The sensitivity of the sensor is

$$s = \frac{dC_f}{dh} = \frac{2\pi\varepsilon_0}{\ln(b/a)}(\varepsilon_r - 1) \quad [\mathrm{F/m}].$$

The sensitivity is governed by the permittivity of the fluid and the dimensions of the two tubes.

Figure 5.13b shows the calculated transfer function of the fuel gauge. In practice the transfer function is slightly nonlinear for very low levels and very high levels of the fluid because of the capacitor edge effects.

Capacitive fuel gauges of this type are often used in diesel fuel tanks on ships and in aircraft fuel tanks. The idea can be used for any fluid that is nonconductive, such as oil or even water.

Capacitive sensors are some of the simplest, most rugged sensors that can be made and are useful in many applications beyond those described here (many more will be encountered in the following chapters). However, with few exceptions the capacitances are small and changes in capacitance even smaller. Therefore they require special methods of transduction. Often, instead of measuring DC voltages, the sensor is made as part of an oscillator whose frequency depends on capacitance and the frequency is measured, usually digitally. Others use an AC source and rather than sensing capacitance, the impedance or phase of the circuit is sensed. We shall discuss some of these issues in **Chapter 11**.

5.3.3 Capacitive Actuators

Capacitive actuation is exceedingly simple and may be understood from **Figure 5.1**. When a potential is connected across the two conductors, they acquire opposite sign charges. These charges attract each other based on Coulomb's law and this force tends to pull the two conductors closer together. Coulomb's law defines the force between the charge and the electric field intensity. Given a charge Q [C] in an electric field intensity $\mathbf{E}$ [V/m], a force $\mathbf{F}$ is exerted on the charge by the electric field as

$$\mathbf{F} = Q\mathbf{E} \quad [\mathrm{N}]. \tag{5.7}$$

The boldface notation indicates that both the electric field intensity and the force are vectors, that is, that they have magnitude and direction in space. In a parallel plate capacitor, the magnitude of the total electric field intensity (i.e., the sum of the electric field produced by the upper and lower charged plates) is given as

$$E = \frac{V}{d} \quad [\mathrm{V/m}], \tag{5.8}$$

where the direction is perpendicular to the plates and the electric field intensity points from the positively charged plate to the negatively charged plate (tending to close the gap between them; that is, they are attracted to each other since the charges on the two plates have opposite signs). Thus mechanical motion of the conductors is possible. In the particular case of the parallel plate capacitor in **Figure 5.2**, the force is found by substituting Q in **Equation (5.7)** with **Equation (5.1)** and E from **Equation (5.8)**, except that the electric field intensity in **Equation (5.8)** must be divided by 2, that is, the force produced by the lower plate on the upper plate is the electric field intensity of the lower plate at the location of the upper plate multiplied by the charge of the upper plate. The capacitance of the parallel plate capacitor is given in **Equation (5.2)**. Putting all these together, the force is

$$F = \frac{CV^2}{2d} = \frac{\varepsilon_0 \varepsilon_r S V^2}{2d^2} \quad [\mathrm{N}]. \tag{5.9}$$

As before, if we cannot assume parallel plates with a small distance d between them, the equation will not be exact, but we can still expect the general relationships to hold—that is, force will be proportional to S, ε, and V^2 and inversely proportional to d^2.

If there is a force, we can also define an energy based on the fact that force is the rate of change of energy over distance:

$$F = \frac{dW}{dl} \quad [\mathrm{N}]. \tag{5.10}$$

Therefore energy is

$$W = \int_0^d \mathbf{F}.d\mathbf{l} = \frac{CV^2}{2} = \frac{\varepsilon_0 \varepsilon_r S V^2}{2d} \quad [\mathrm{J}]. \tag{5.11}$$

Note that this can also be written as

$$W = \frac{\varepsilon(Sd)}{2}\left(\frac{V^2}{d^2}\right) = \frac{\varepsilon E^2}{2} v \quad [\mathrm{J}], \tag{5.12}$$

where v is the volume of the space between the capacitor's plates, ε is the permittivity of the medium, and $E = V/d$ is the electric field intensity between the plates. The quantity $\varepsilon E^2/2$ has units of [J/m^3] and is therefore the energy density in the capacitor.

Given a fixed conductor, a second conductor, if connected to a potential difference with respect to the first, will move with respect to the fixed conductor. This is shown schematically in **Figure 5.14**. This motion may be used for positioning or, as is the case here, as an electrostatic loudspeaker.

However, inspection of **Equation (5.9)** shows that the forces that can be achieved are rather small since ε_0 is very small. To make this device perform useful work, one of two things must be accomplished: either make the distance very small (very limited

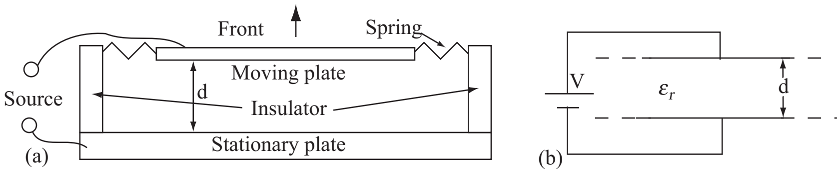


FIGURE 5.14 ■ (a) Schematic structure of a capacitive actuator and (b) its equivalent circuit.

motion) or increase the voltage to high values. In microelectromechanical (MEM) actuators, the distances are naturally small, but the voltage is also small (otherwise electric breakdown may occur). In electrostatic loudspeakers and headphones the displacement must be large so a high voltage (up to a few thousand volts) must be used.

But the obvious vertical attraction between two plates of a parallel plate capacitor is not the only mechanism for electrostatic actuation. Consider the asymmetric configuration of the two plates of a parallel plate capacitor in **Figure 5.15a**. The two plates attract each other, but now the force has a vertical and a horizontal component. The plates not only will attract and try to close the distance between them, but also to close the horizontal separation. This basic method may also be used to affect rotational motion using the device in **Figure 5.15b**. Here the distance between the plates is fixed, but the plates can rotate with respect to each other and will tend to do so until they are positioned exactly overlapping each other. With the addition of a spring to restore the initial position, this can be made into a very accurate positioner (see **Example 5.4**). The forces involved can be calculated from the virtual displacement principle, as follows: Using again the plates in **Figure 5.15a**, suppose we allow the upper plate to move a "virtual" distance dx to the left. The volume between the plates changes by a quantity $dv = wddx$, where w is the width of the plate and d is the distance between them. The change in energy due to this motion is dW and necessarily this must be equal to Fdx. With the energy density defined as $\varepsilon E^2/2$, we write

$$Fdx = dW = \frac{\varepsilon E^2}{2} wddx \tag{5.13}$$

and the magnitude of the lateral force is

$$F = \frac{\varepsilon E^2}{2} wd \quad [\text{N}]. \tag{5.14}$$

The direction of this force is such as to force the plates to the center with respect to each other since that is a situation of minimum (zero) lateral force.

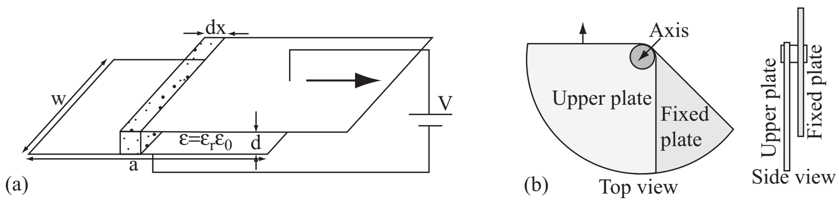


FIGURE 5.15 ■ Capacitive actuators. (a) Linear actuator. The lower plate, separated a distance d from the upper plate, is fixed. The upper plate can move relative to the lower plate. (b) Rotary actuator. The upper plate rotates under the influence of charges on the plates.

EXAMPLE 5.3 Electrostatic air pump

A small air pump is made as shown in **Figure 5.16a**. The body of the pump is a cylindrical cavity of radius $a = 50$ mm and height $h = 1$ mm. The top part is made of a disk of radius $b = 40$ mm suspended on a rubber diaphragm to allow the plate to move up and down. An insulating stop ring, 0.1 mm thick, prevents the moving disk from shorting to the body during compression. Two flap valves are used to affect pumping. The lower valve prevents air from leaking during the compression stroke, whereas the side valve closes during the suction stroke so air enters through the lower flap. Assume the rubber diaphragm is sufficiently flexible to offer no resistance so that the moving plate can move down until it encounters the plastic ring on the bottom plate. Assuming as well that the moving and stationary plates form a parallel plate capacitor and that the mass of the moving plate can be neglected, calculate

a. The volume of air the pump can move per minute if the source is sinusoidal at 60 Hz.
b. The maximum pressure at the outlet port if the peak voltage is 200 V.

Solution: The moving and stationary plates attract each other during each half cycle of the AC signal since during the positive half cycle the moving plate is positive with respect to the stationary plate, whereas during the negative half cycle the moving plate is negative with respect to the stationary body. Therefore the pump will effectuate 120 strokes/s. In each stroke the downward motion of the plate expels a volume of air equal to the change in volume of the chamber (**Figure 5.16b**). The maximum pressure occurs throughout the chamber when the plate rests against the plastic ring and is equal to the attraction force divided by the area of the moving plate.

a. The volume expelled per stroke is the volume of the chamber in **Figure 5.16b**. We write for this volume

$$\begin{aligned}\Delta v &= \pi a^2 h - \pi a^2 c - \pi\left(a^2 - b^2\right)\frac{h-c}{2} \\ &= \pi \times 50^2 \times 1 - \pi \times 50^2 \times 0.1 - \pi\left(50^2 - 40^2\right)\frac{1-0.1}{2} \\ &= \pi(2500 - 250 - 405) = 5796.24\ \text{mm}^3\end{aligned}$$

Since there are 120 strokes/s, or 120×60 strokes/min, the volume of air expelled per minute is

$$v = 120 \times 60\Delta v = 120 \times 60 \times 5796.24 = 4.17 \times 10^7\ \text{mm}^3.$$

The maximum volume that the pump can move is 0.0417 m^3/min or 41.7 L/min.

b. The force between the moving plate and the body of the pump is given in **Equation (5.9)**:

$$F = \frac{\varepsilon_0 S V^2}{2d^2} \quad [\text{N}]$$

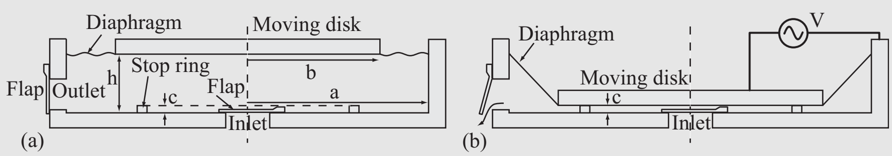


FIGURE 5.16 ■ (a) Principle of an electrostatic pump before application of the potential. (b) The electrostatic pump at the end of the compression stroke.

where ε_0 is the permittivity of air, S is the area of the moving plate, and d is the distance between the moving plate and the bottom surface. This varies from a minimum when $d = h = 1$ mm and a maximum when $d = c = 0.1$ mm. The maximum force is therefore

$$F = \frac{\varepsilon_0 S V^2}{2d^2} = \frac{8.854 \times 10^{-12} \times \pi \times 0.04^2 \times 200^2}{2 \times 0.0001^2} = 0.089 \text{ N}.$$

The pressure is

$$P = \frac{F}{S} = \frac{0.089}{\pi \times 0.04^2} = 17.7 \text{ N/m}^2.$$

This is a mere 17.7 Pa, a very low pressure indicating that the forces involved are very small, something that is typical in electrostatic devices. In fact, because of friction and tension in the diaphragm and flaps, this pressure may be even lower. Nevertheless, very small electrostatic forces are useful in microdevices, as we shall see in **Chapter 10**, where the dimensions and distances are much smaller than in this example.

EXAMPLE 5.4 Rotary capacitive actuator

Consider the rotary capacitive actuator in **Figure 5.15b**. The actuator is made of two half-circle plates of radius $a = 5$ cm and separated a distance $d = 0.5$ mm with a sheet of plastic of permittivity $\varepsilon = 4\varepsilon_0$ to separate the plates. Given the dimensions and permittivity of air and assuming the distance between the plates to be small enough to justify the use of formulas for the parallel plate capacitor, calculate (neglect any friction between the moving plate and plastic)

a. The force that the moving plate exerts as a function of the applied voltage.

b. The torque the moving plate can supply as a function of the applied voltage.

Solution:
a. Using the process in **Equations (5.13)** and **(5.14)** and referring to **Figure 5.17**, we calculate the force as follows: Suppose the lower plate rotates an angle $d\theta$. The change in volume is

$$dv = \frac{(ad\theta)ad}{2} \quad [\text{m}^3],$$

where the quantity $ad\theta$ is the arc length shown and the dotted surface area is taken as a triangle. The change in energy due to this (virtual) motion is

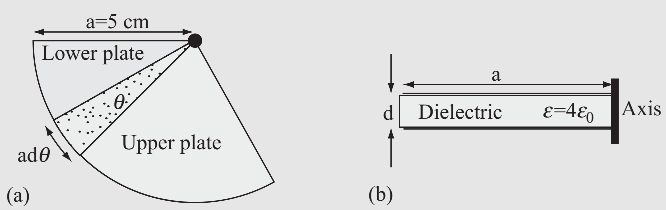


FIGURE 5.17 ■ The rotary capacitive actuator. (a) Top view. (b) Side view.

$$dW = \varepsilon \frac{E^2}{2} dv = \varepsilon \frac{V^2}{4d^2} a^2 d d\theta = \frac{\varepsilon V^2 a^2}{4d} d\theta \quad [\mathrm{J}].$$

By definition, force is the rate of change in energy:

$$F = \frac{dW}{d\theta} = \frac{\varepsilon V^2 a^2}{4d} \quad [\mathrm{N}].$$

The force acts at the center of gravity of the disk and depends on the applied voltage, permittivity between the plates, distance between the plates, and the radius of the plate.

Numerically we get

$$F = \frac{4 \times 8.845 \times 10^{-12} \times (0.05)^2}{4 \times 0.0005} V^2 = 44.27 \times 10^{-12} V^2 \ [\mathrm{N}].$$

b. The torque is the product of force and radial distance. Since the force acts at the center of gravity we must calculate its location or, at the very least, estimate it. In this case the result can be found in tables. The center of gravity is found at a radial distance of $4a\sqrt{2}/3\pi$ from the axis on the line dividing the quarter-circle plate. The torque is therefore

$$T = Fl = \frac{\varepsilon V^2 a^2}{4d} \frac{4a\sqrt{2}}{3\pi} = \frac{\varepsilon V^2 a^3 \sqrt{2}}{3\pi d} \quad [\mathrm{N} \cdot \mathrm{m}].$$

Or numerically,

$$T = \frac{\varepsilon V^2 a^3 \sqrt{2}}{3\pi d} = \frac{4 \times 8.845 \times 10^{-12} \times (0.05)^3 \sqrt{2}}{3 \times \pi \times 0.0005} V^2 = 1.33 \times 10^{-12} V^2 \ [\mathrm{N} \cdot \mathrm{m}].$$

As expected, the force and torque are very small but increase quickly with the applied voltage. Although this actuator cannot be expected to be of use in regular applications, it is practical in MEMS devices, to be discussed in **Chapter 10**.

5.4 | MAGNETIC FIELDS: SENSORS AND ACTUATORS

Magnetic sensors and actuators are those governed by the magnetic field (more specifically, by the magnetic flux density, **B**) and its effects. The magnetic flux density is also called magnetic induction. Therefore these sensors tend to be inductive sensors. We should be careful, however, because induction has other connotations that will become clear later on. As we have done with electric sensors and actuators, we will try to rely as much as possible on simple properties of the magnetic field without getting into the nitty-gritty aspects of the theory (which would require a full understanding of Maxwell's equations). Therefore much of the work will rely on inductance, magnetic circuits, and magnetic forces, which can be explained, at least qualitatively, without resorting to Maxwell's equations. Needless to say, because of that, some quantities will be approximate and some merely qualitative.

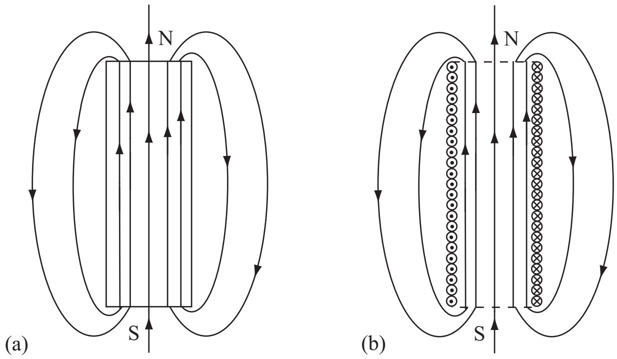


FIGURE 5.18 ■ (a) Permanent magnet. (b) Coil with field equivalent to that of the magnet in (a).

To start with, a magnetic field may be understood by resorting to a permanent magnet. The magnet exerts a force on another magnet through space. We say that a "field" exists around the magnet through which it interacts. This force field is in fact the magnetic field (**Figure 5.18a**). The same can be observed by driving a current through a coil (**Figure 5.18b**). Since the two fields in **Figure 5.18** are identical, their sources must be identical and we conclude that all magnetic fields are generated by currents. In the case of the permanent magnet, the currents are atomic currents produced by electrons. A magnet will attract or repel another magnet—this gives us the first observable interaction in the magnetic field—but it will also attract a piece of iron. On the other hand, it will not attract a piece of copper. The conclusion is that there are different types of materials in terms of their magnetic properties. These properties are governed by the permeability of the material, μ [H/m]. The "strength" of the magnetic field is usually given by the magnetic flux density, **B** [T], or the magnetic field intensity, **H** [A/m]. The relation between the two fields is

$$\mathbf{B} = \mu_0 \mu_r \mathbf{H} \quad [\mathrm{T}], \tag{5.15}$$

where $\mu_0 = 4\pi \times 10^{-7}$ H/m is the permeability of vacuum and μ_r is the relative permeability of the medium in which the relation holds. μ_r is the ratio between the permeability of the medium and that of vacuum and hence is a dimensionless quantity associated with each material in nature. **Tables 5.3a–5.3d** list the permeabilities of some useful materials and classify them based on their relative permeabilities. If the relative permeability is <1, the materials are called diamagnetic, if >1, paramagnetic. Materials with $\mu_r >> 1$ are called ferromagnetic (ironlike) and are often the most useful materials

TABLE 5.3a ■ Permeability of various diamagnetic and paramagnetic materials

Material	Relative permeability	Material	Relative permeability
Silver	0.999974	Air	1.00000036
Water	0.9999991	Aluminum	1.000021
Copper	0.999991	Palladium	1.0008
Mercury	0.999968	Platinum	1.00029
Lead	0.999983	Tungsten	1.000068
Gold	0.999998	Magnesium	1.00000693
Graphite (Carbon)	0.999956	Manganese	1.000125
Hydrogen	0.999999998	Oxygen	1.0000019

TABLE 5.3b ■ Permeability of various ferromagnetic materials

Material	μ_r	Material	μ_r
Cobalt	250	Permalloy (78.5% Ni)	100,000
Nickel	600	Fe_3O_4 (magnetite)	100
Iron	6000	Ferrites	5000
Supermalloy (5% Mo, 79% Ni)	10^7	Mumetal (75% Ni, 5% Cu, 2% Cr)	100,000
Steel (0.9% C)	100	Permendur	5000
Silicon iron (4% Si)	7000		

TABLE 5.3c ■ Permeability of various soft magnetic materials

Material	Relative permeability (maximum) μ_r
Iron (0.2% impure)	9000
Pure iron (0.05% impure)	2×10^5
Silicon iron (3% Si)	55,000
Permalloy	10^6
Supermalloy (5% Mo, 79% Ni)	10^7
Permendur	5000
Nickel	600

TABLE 5.3d ■ Permeability of various hard magnetic materials

Material	μ_r
Alnico (aluminum-nickel-cobalt)	3–5
Ferrite (barium-iron)	1.1
Sm-Co (samarium-cobalt)	1.05
Ne-Fe-B (neodymium-iron-boron)	1.05

when working with magnetic fields. There are other types of magnetic materials (ferrites, magnetic powders, magnetic fluids, magnetic glasses, etc.) that we will encounter, and will describe them as necessary. Soft magnetic materials are those in which magnetization is reversible (i.e., they do not become permanent magnets following application of an external magnetic field), whereas hard magnetic materials are materials that retain magnetization and therefore are often used for production of permanent magnets.

Magnetic materials, especially ferromagnetic materials, have two related and important properties in addition to those discussed above. One is magnetic hysteresis and the other is nonlinearity of its magnetization curve. Magnetic hysteresis is shown in **Figure 5.19**. It indicates that as the magnetization changes (as shown by the change in magnetic field intensity), the curve traces different paths as the magnetization increases and as it decreases. The area of the magnetization curve is associated with losses. From a sensing point of view it is important to understand that the narrower the curve, the easier it is to reverse magnetization. This indicates that these materials are appropriate as magnetic cores in structures such as electric motors or transformers, especially those operating on alternating currents. Wide magnetization curves mean that the magnetization is not easily reversed, and these are usually materials used in permanent magnets.

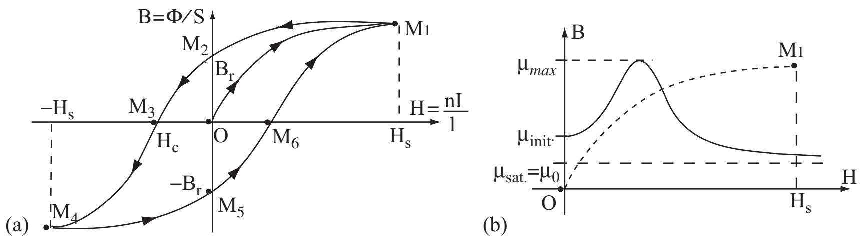


FIGURE 5.19 ■ Hysteresis (magnetization) curve and the resulting permeability for a ferromagnetic material.

Permeability is the slope of the magnetization curve (**Figure 5.19b**). Since this slope changes from point to point, permeability is nonlinear.

An important relation is that between current and magnetic flux density. Consider a very long straight wire (infinitely long) carrying a current I and placed in a medium of permeability $\mu = \mu_0\mu_r$. The magnitude of the magnetic flux density is

$$B = \mu_0\mu_r \frac{I}{2\pi r} \quad [\mathrm{T}], \tag{5.16}$$

where r is the distance from the wire to the location where the field is calculated (**Figure 5.20a**). The magnetic field has a direction as shown. If we place a system of coordinates (cylindrical in this case), the field can be described as a vector:

$$\mathbf{B} = \hat{\boldsymbol{\phi}}\mu_0\mu_r \frac{I}{2\pi r} \quad [\mathrm{T}]. \tag{5.17}$$

The important point to observe is that the field is perpendicular to the current. The relation between current and field is given by the right-hand rule shown in **Figure 5.20b**.

In more practical configurations, the wire may not be very long or it may be wound in a coil, but, nevertheless, the basic relations hold: the larger the current and/or the permeability, or the shorter the distance between the current and the location where the magnetic field is needed, the larger the magnetic flux density. However, **Equation (5.16)** is only correct for long, thin wires. In other configurations, the flux density may be quite different. In a long solenoid, with n turns per unit length (see **Figure 5.21a**), the magnetic flux density in the solenoid is constant, and with the geometry and system of coordinates shown is

$$\mathbf{B} = \hat{\mathbf{z}}\mu_0\mu_r nI \quad [\mathrm{T}]. \tag{5.18}$$

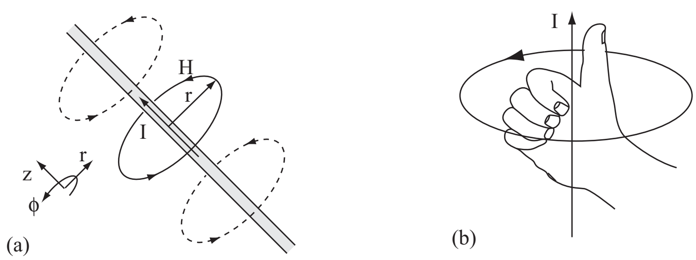


FIGURE 5.20 ■ (a) Magnetic field of a long, straight current-carrying conductor. (b) The relation between the direction of current and field (right-hand rule).

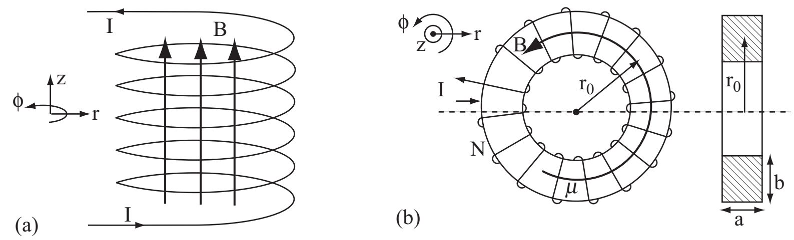


FIGURE 5.21 ■ (a) The long solenoid and its field. (b) The toroidal coil and its field.

The magnetic flux density outside the solenoid is zero. Similarly, the magnetic flux density in a toroidal coil of average radius r_0, cross-sectional area S, with N turns uniformly wound on the torus (**Figure 5.21b**) is

$$\mathbf{B} = \hat{\mathbf{z}}\frac{\mu_0\mu_r nI}{2\pi r_0} \quad [\mathrm{T}]. \tag{5.19}$$

The flux density outside the toroidal coil is zero. In other configurations the magnetic flux density is more complex and may not be calculable analytically.

If the flux density is integrated over an area, we obtain the flux of the magnetic field over that area:

$$\Phi = \int_S \mathbf{B}\cdot d\mathbf{s} \quad [\mathrm{Wb}]. \tag{5.20}$$

Of course, if **B** is constant over an area S and at an angle θ to the surface, the flux is $\Phi = BS\cos\theta$, as indicated by the scalar product in **Equation (5.20)**.

The force in a magnetic field is based on the fact that a charge moving at a velocity **v** in a magnetic field **B** experiences a force, called the **Lorentz force**, given as

$$\mathbf{F} = q\mathbf{v}\times\mathbf{B} \quad [\mathrm{N}], \tag{5.21}$$

where the force is perpendicular to the direction of both **v** and **B**. The magnitude of the Lorentz force may be written as

$$F = qvB\sin\theta_{vB} \quad [\mathrm{N}], \tag{5.22}$$

where θ_{vB} is the angle between the direction of motion of the charge q and the direction of **B** as shown in **Figure 5.22a**. In most sensing applications charges are not moving in space (in some cases they do), but rather in conductors. In these important cases the force may be recast to act on the current rather than on charges. The starting relation is

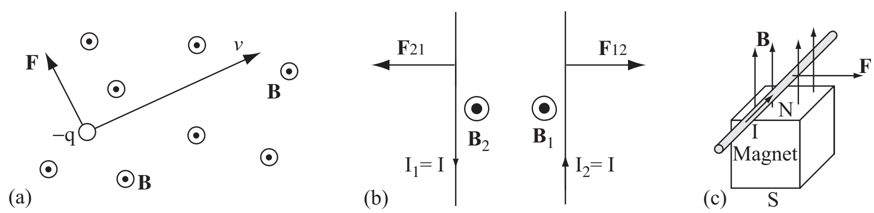


FIGURE 5.22 ■ (a) The relation between force and magnetic field for a moving charge. (b) Forces exerted by oppositely flowing currents on each other. (c) Force exerted by a magnet on a current-carrying wire.

Equation (5.21) and the current density in a volume containing n charges per unit volume is

$$\mathbf{J} = nq\mathbf{v} \quad [\mathrm{A/m^2}]. \tag{5.23}$$

That is, the force in **Equation (5.21)** can be written as a force per unit volume,

$$\mathbf{f} = \mathbf{J} \times \mathbf{B} \quad [\mathrm{N/m^3}], \tag{5.24}$$

or it may be integrated over a volume in which the current density flows to calculate the total force on the current:

$$\mathbf{F} = \int_v \mathbf{J} \times \mathbf{B} dv \quad [\mathrm{N}]. \tag{5.25}$$

As previously, the magnitude of the force density may be written as $f = JB\sin\theta$, where θ is the angle between the flux density and the current density.

The force between two wires carrying currents in opposite directions is shown in **Figure 5.22b**. The forces the wires exert on each other are in opposite directions and tend to separate the wires. These forces are evaluated from **Equations (5.17)** and **(5.25)**. If the currents were in the same direction (or the magnetic field reversed) the wires would attract. For long parallel wires, the force for a length L of the wire is

$$F = BIL \quad [\mathrm{N}]. \tag{5.26}$$

For other configurations the relation is much more complicated, but, in general, force is proportional to B, I, and L. A single wire carrying a current will be attracted or repelled by a permanent magnet, as shown in **Figure 5.22c**. These principles are the basis for magnetic actuation, and unlike electric field actuators, the forces can be very large since B, I, and L can be controlled and can be quite large. The relations above, and some we will use below, are very simple, but they suffice to help us understand the behavior of magnetic devices, at least qualitatively, and to explain how sensors and actuators operate.

5.4.1 Inductive Sensors

Inductance is a property of a magnetic device in the manner that capacitance is a property of an electric device. Inductance is defined as the ratio of magnetic flux and current:

$$L = \frac{\Phi}{I} \, \frac{\text{webber}}{\text{ampere}} \text{ or [henry]}. \tag{5.27}$$

The unit of inductance is the henry (H). Inductance is independent of current since $\mathbf{\Phi}$ is linearly dependent on current (see **Equations (5.17)** and **(5.20)**). All magnetic devices have an inductance, but inductance is most often associated with electromagnets—in which the flux is produced by a current through conductors, usually in the form of coils. We define two types of inductance:

1. Self inductance: the ratio of the flux produced by a circuit (a conductor or a coil) in itself and the current that produces it. That is, the flux in **Equation (5.20)** is the flux through the device itself. Usually denoted as L_{ii}.

2. Mutual inductance: the ratio of the flux produced by circuit i in circuit j and the current in circuit i that produced it. Usually denoted as M_{ij}.

The concept of self-inductance is shown in **Figure 5.23a**, that of mutual inductance in **Figure 5.23b**. Thus any circuit (conductor, coil of conductors) has a self-inductance. A mutual inductance exists between any two circuits as long as there is a magnetic field (flux) that couples the two when a current passes through either circuit. This coupling can be large (tightly coupled circuits) or small (loosely coupled circuits).

Whereas the measurement of inductance is relatively easy, the calculation of inductance is not and depends on the geometry and its details. Nevertheless, exact or approximate formulas for inductance of various coils exist. For example, for a long circular coil of radius r and with n turns/m, the self-inductance can be approximated as

$$L = \mu n^2 \pi r^2 \quad [\mathrm{H/m}]. \tag{5.28}$$

The self-inductance of toroidal coils can also be calculated relatively easily. Other approximate formulas for short coils as well as for inductance of straight wires are available.

It should also be recalled that the relation between voltage on an inductor and the current through it is strictly an AC relation given as

$$V = L\frac{dI}{dt} \quad [\mathrm{V}], \tag{5.29}$$

where $I(t)$ is the current in the inductor and L is its total inductance. This voltage is called back-emf because its polarity opposes the polarity of the source that provides the current.

If a loop or coil that contains N turns is placed in the field produced by a second coil, the induced voltage (often called induced electromotive force or emf) is

$$emf = -N\frac{d\Phi}{dt} \quad [\mathrm{V}], \tag{5.30}$$

where the negative sign indicates a phase difference between the current producing the flux and the voltage induced in the coil. For example, using **Figure 5.23b**, the voltage induced in coil 2 due to the field produced by coil 1 would be $-N_2 d\Phi_{12}/dt$, where N_2 is the number of turns in the coil on the right and Φ_{12} is the flux produced by the coil on the left in the coil on the right. Clearly **Equation (5.30)** is general and its usefulness depends on our ability to evaluate fluxes. In some cases that is relatively easy, but in others it is not.

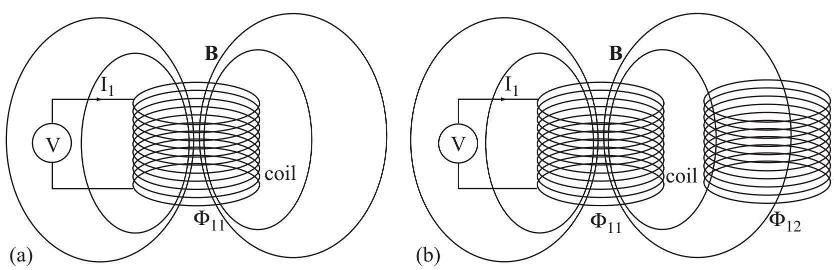


FIGURE 5.23 ■ The concept of inductance. (a) Self-inductance. (b) Self- and mutual inductance. Mutual inductance is due to the flux linking the two coils.

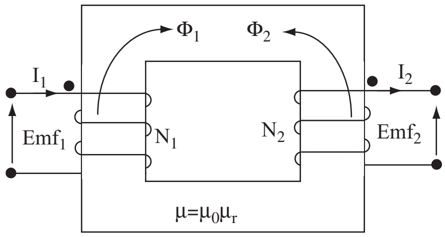


FIGURE 5.24 ■ The transformer.

The concepts of self- and mutual inductance are also associated with the principle of the transformer. In a transformer, which will contain two or more coils, an AC voltage applied to one circuit (or coil) produces a voltage in any other circuit that couples to the driving coil, as shown in **Figure 5.24**. The coils, with N_1 and N_2 turns, respectively, produce fluxes when a current exists through them. All flux produced by coil 1 couples with coil 2 through the magnetic circuit made of a ferromagnetic material (iron, for example). The voltages and currents relate as follows:

$$V_2 = \frac{N_2}{N_1} V_1 = \frac{1}{a} V_1, \; I_2 = \frac{N_1}{N_2} I_1 = aI_1, \tag{5.31}$$

where $a = N_1/N_2$ is the transformer ratio. If not all the flux produced by one coil links the other, the ratios above must be multiplied by a coupling coefficient that indicates how tightly the two coils are coupled. Under this condition the device is a loosely coupled transformer and occurs when the core is made of low-permeability materials such as air or when the core is not closed.

Most inductive sensors rely on self-inductance, mutual inductance, or transformer concepts to operate. It should be remembered, however, that these are active elements and require connection to a source to produce an output. In the process, the magnetic field described above is produced and the sensor can be said to respond to changes in this magnetic field. The output is sometimes given in terms of the magnetic flux density, as are sensitivity and error, but more often in terms of the output voltage the sensor produces. Similar considerations apply to actuators.

The most common types of stimuli sensed by inductive sensors are position (proximity), displacement, and material composition. These will be described next, taking advantage of the concepts of inductance and magnetic circuits. There are also sensors that use inductance and induction indirectly and some of these will be encountered later in this chapter and in later chapters.

5.4.1.1 Inductive Proximity Sensors

An inductive proximity sensor contains, at the very least, a coil (inductor) that, when a current passes through it, generates a magnetic field, as shown in **Figure 5.25a**. The coil has an inductance, which depends on the dimensions of the coil, number of turns, and materials around it. The current and the diameter of the coil define the extent to which the field projects away from the coil and therefore the range and span of the sensor. As the sensor gets closer to the sensed surface (**Figure 5.25b**), the inductance of the coil increases if the sensed surface is ferromagnetic (it does not change or changes very little if the surface is not ferromagnetic, but we shall see shortly that the same can be achieved in nonferromagnetic materials, provided they are conducting and the field is alternating). It is then sufficient to use an "inductance meter" and a calibration curve (transfer

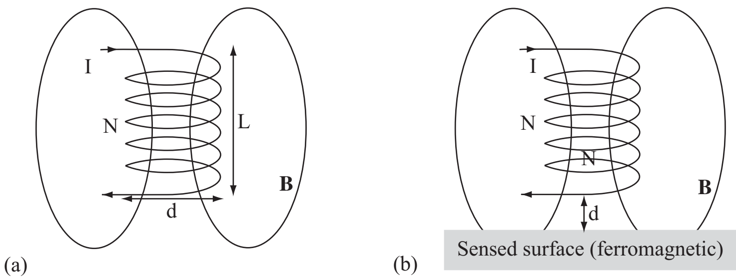


FIGURE 5.25 ■ The basic inductive proximity sensor. (a) Coil in air (no sensed surface). (b) The field produced by the coil interacts with the sensed surface, changing its inductance.

function) to devise a sensor. An inductance meter is usually made of an AC current source and a voltmeter or an AC bridge. By measuring the voltage across the inductor, the impedance can be evaluated, and since $Z = R + j\omega L$ (R is the ohmic resistance and $\omega = 2\pi f$ is the angular frequency), the inductance L is immediately available as a measure of the position of the coil or proximity to the surface being sensed. We assumed here that R is constant, but even if it is not, the impedance in air (nothing being sensed) is known and this can be used for calibration.

The advantage of the device in **Figure 5.25** is its simplicity, but it is a highly nonlinear sensor, and in practical sensors a few things are done to improve its performance. First, a ferromagnetic core is added to increase the inductance of the sensor. Most often this is made of a ferrite material (a powdered magnetic material such as iron oxide [Fe_2O_3] or other oxides in a binding substance and sintered into the shape needed). Ferrites have the advantage of being high-resistance materials (low conductivity). In addition, a shield may be placed around the sensor to prevent sensitivity to objects on the side of the sensor or at its back (**Figure 5.26a,b**). The net effect of the shield is to project the field in front of the sensor and hence increase both the field (inductance) and the span of the sensor. In other sensors there are two coils, one serving as a reference and the other as a sensor (**Figure 5.26b**). The reference coil's inductance remains constant and the two are balanced. When a surface is sensed, the sensing coil's inductance increases and the imbalance between the coils serves as a measure of distance (differential sensor). Other sensors, like the one is **Figure 5.26c**, may employ a closed magnetic circuit, which tends to concentrate the magnetic field in the gaps, and usually do not require shielding since the magnetic field is constrained within the ferromagnetic material making up the core. In all cases, exact calculation of the fields and the sensor response is only possible using numerical tools. In many cases, however, these can be determined rather easily by experiment.

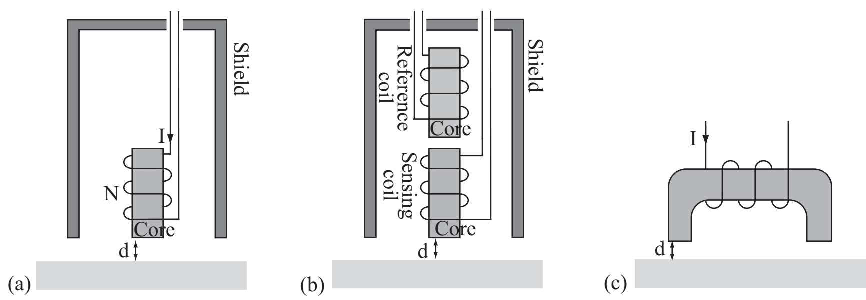


FIGURE 5.26 ■ Practical proximity sensors. (a) Shielded sensor. (b) Shielded sensor with reference coil. (c) Sensor employing a magnetic circuit to concentrate the field in a small gap.

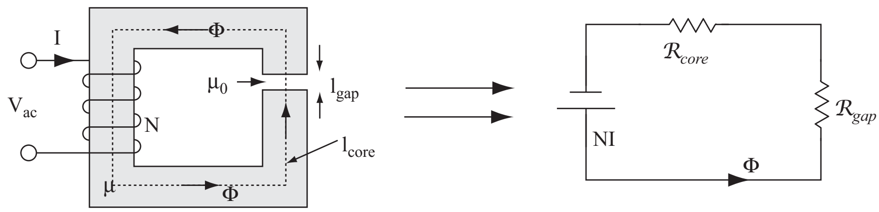


FIGURE 5.27 ■ The concept of a magnetic circuit and its equivalent as an electric circuit with modified quantities.

In certain cases, especially in transformers and in closed magnetic circuits such as the sensor shown in **Figure 5.26c**, an approximate method can be used to calculate parameters such as inductance, magnetic flux, or induced voltage based on the concept of the ***magnetic circuit***. In this method, the magnetic flux is viewed as a "current," the term NI (i.e., the product of the current) and the number of turns in a coil as a "voltage," and a term called reluctance replaces the resistance in a regular circuit. The basic concept is shown in **Figure 5.27**. The reluctance of a member of the magnetic circuit of length l_m [m], permeability μ [H/m], and cross-sectional area S [m^2] is

$$\mathcal{R}_m = \frac{l_m}{\mu S} \quad [1/\mathrm{H}]. \tag{5.32}$$

With these preliminaries, the flux in the circuit is

$$\Phi = \frac{\sum NI}{\sum \mathcal{R}_m} \qquad [\mathrm{Wb}]. \tag{5.33}$$

In the case shown in **Figure 5.27**, there are two reluctances, one due to the path in the core (l_{core}) and one due to the path in the gap (l_{gap}).

Other quantities such as induced voltages in coils, forces, and fields can be calculated from these relations.

However, the equivalent circuit reveals the requirements as well. The flux must "flow" in the closed circuit just like an electric current must flow in a closed circuit. This can be achieved, at least approximately, if the permeability of the core is high. When the permeability is low, such as in the gap, the length of the gap must be short to prevent flux from "spreading" out and invalidating the assumption of the circuit. Nevertheless, the method is useful for approximate calculations. Magnetic circuits can be used with DC or AC sources.

EXAMPLE 5.5 Paint thickness sensor

A paint thickness sensor is built as in **Figure 5.26c** with a magnetic field sensor (a sensor that can quantify the magnetic flux density—a Hall element—to be discussed in **Section 5.4.2**) embedded into one of the surfaces as shown in **Figure 5.28**. The sensor's core is made of silicon steel (a steel often employed in electromagnetic devices because it has high permeability and low conductivity). The coil includes $N = 600$ turns and is supplied with a DC current $I = 0.1$ A. The core of the sensor has a cross-sectional area $S = 1$ cm^2 and an average magnetic path length $l_c = 5$ cm. The sensor can only be used reliably to test for paint thickness on ferromagnetic materials such as steel, cast iron, nickel, and their alloys. It can also sense the thickness of plating layers such as zinc or copper on steel. The purpose here is to establish a transfer function and sensitivity for the

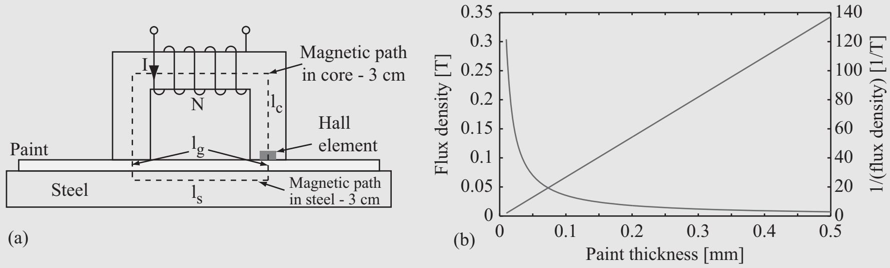


FIGURE 5.28 ■ Paint thickness sensor. (a) Geometry and dimensions. (b) Flux density (B) measured as a function of paint thickness and the reciprocal of flux density (1/*B*) as a function of paint thickness.

sensor for paint thicknesses from 0.01 mm to 0.5 mm. The measurand is the paint thickness and the measured quantity is the magnetic flux density as measured by the field sensor. Silicon steel has a relative permeability of 5000, whereas steel has a relative permeability of 1000. Paint is nonmagnetic with relative permeability of 1.

Solution: We start by calculating the reluctances of the core, the two gaps, and an estimate of the reluctance in the steel material. Then we calculate the flux and finally the flux density as a function of gap thickness.

The reluctances of the core and the gap are

$$\mathcal{R}_c = \frac{l_c}{\mu_c S} = \frac{5 \times 10^{-2}}{5000 \times 4\pi \times 10^{-7} \times 10^{-4}} = 7.957 \times 10^4/\text{H}$$

and

$$\mathcal{R}_g = \frac{l_g}{\mu_c S} = \frac{l_g}{4\pi \times 10^{-7} \times 10^{-4}} = 7.957 \times 10^9 l_g/\text{H}.$$

Note that the reluctance of the gap is much higher than that of the core because of the low permeability of air.

In the steel we have the length of the path, but not the cross-sectional area. As a first approximation we will assume the same cross-sectional area as the core. In practice this unknown quantity will be part of the calibration of the sensor. With this assumption we have, for steel,

$$\mathcal{R}_s = \frac{l_s}{\mu_c S} = \frac{0.03}{1000 \times 4\pi \times 10^{-7} \times 10^{-4}} = 2.387 \times 10^5/\text{H}.$$

Now we can calculate the flux in the core, gap, and steel:

$$\Phi = \frac{NI}{\mathcal{R}_c + 2\mathcal{R}_g + \mathcal{R}_s} = \frac{600 \times 0.1}{7.957 \times 10^4 + 2 \times 7.957 \times 10^9 l_g + 2.387 \times 10^5}$$
$$= \frac{60}{3.183 \times 10^5 + 2 \times 7.957 \times 10^9 l_g} \quad [\text{Wb}].$$

Note that because the reluctance of the gap is at least four orders of magnitude larger than that of steel or of the core, the reluctances of steel and the core can be neglected and we obtain the approximate flux above. The flux density is the flux divided by the cross-sectional area:

$$B = \frac{\Phi}{S} = \frac{1}{1 \times 10^{-4}} \left(\frac{60}{3.183 \times 10^5 + 2 \times 7.957 \times 10^9 l_g} \right) = \frac{6}{3.183 + 2 \times 7.957 \times 10^4 l_g} \quad [\text{T}].$$

We could have neglected the first term in the denominator as small, but that would not be the case for very low values of l_g. The transfer function can now be established by entering values of the gap from $l_g = 0.01$ mm to 0.5 mm. The transfer function is shown in **Figure 5.28b**. The curve is highly nonlinear, but nevertheless, the relation between flux density and paint thickness is distinct and usable. In an instrument one can invert the result, that is, as a postprocessing step one can calculate the quantity $1/B$ and plot it against the paint thickness, also shown in **Figure 5.28b**. The result is a linear curve and a much easier to read output. The output units can be calibrated directly in terms of paint thickness.

5.4.1.2 Eddy Current Proximity Sensors

Inductive proximity sensors are sensitive to the presence (proximity) of nonconducting ferromagnetic materials or to any conducting media. Nonconductors in general do not affect inductive proximity sensors. Many inductive proximity sensors are of the eddy current type. The name eddy current comes from the fundamental property of AC magnetic fields to induce currents in conducting media (ferromagnetic or not). This is shown schematically in **Figure 5.29**. There are two related phenomena at work. The currents produced in the conductor, called eddy currents because they flow in closed loops, cause a field that opposes the original field that produces them (Lenz's law). This field reduces the net flux through the sensor's coil. Second, the currents flowing in the conductor being sensed dissipate power. The sensing coil is now forced to supply more power than it would otherwise supply and hence, given a constant current, its effective resistance increases. This change in impedance from $Z = R + j\omega L$ to $Z' = R' + j\omega L'$ is easily sensed either in absolute terms or as a change in the phase of the measured voltage (given a constant current). An AC magnetic field penetrating into a conducting medium is attenuated exponentially from the surface inward (and so are the eddy currents and other related quantities):

$$B = B_0 e^{-d/\delta} \quad [\text{T}], \text{or } J = J_0 e^{-d/\delta} \quad [\text{A/m}^2], \tag{5.34}$$

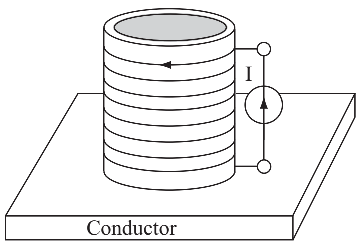


FIGURE 5.29 ■ An eddy current sensor. The AC in the coil induces eddy currents in the conducting plate.

where B_0 and J_0 are the flux density and the eddy current density at the surface, d is the depth in the medium, and δ is the skin depth. Skin depth is defined as the depth at which the field (or current density) is attenuated to $1/e$ of its value at the surface and for planar conductors is given as

$$\delta = \frac{1}{\sqrt{\pi f \mu \sigma}} \quad [\text{m}], \tag{5.35}$$

where f is the frequency of the field. Clearly then, penetration depends on frequency, conductivity, and permeability. The main implication here is that the sensed conductor must be thick enough compared to skin depth. Alternatively, operation at higher frequencies may be needed to reduce the skin depth.

Figure 5.30 shows a number of inductive proximity sensors used for industrial control. A proximity sensor (either capacitive or inductive) can be used to sense distance. However, except for very short distances, their transfer function is much too nonlinear and their span too small to be effective for this purpose. For this reason proximity sensors are often used as switches to provide a clear indication when a certain, preset distance is reached. As with capacitive sensors, inductive sensors can produce an electric output, such as voltage, based on the change in their impedance, but often the inductor is part of an oscillator (LC oscillator is the most common, in which case $f = 1/2\pi\sqrt{LC}$) and the frequency of the sensor is then used as the output.

Figure 5.31 shows two eddy current sensors used for nondestructive testing of materials. The sensor on top is an absolute EC sensor (or probe) since it contains a single coil and measures the absolute change in impedance due to the presence of small flaws in conducting materials. The bottom sensor contains two small coils separated a short distance apart and is used as a differential sensor (see **Example 5.6**). The sensor on top operates at 100 kHz and its coil is embedded in a dielectric. The bottom sensor operates at 400 kHz and its coils are embedded in ferrite. **Figure 5.32** shows two differential eddy current probes as used for flaw detection inside tubes. The top sensor is 19 mm in diameter and is designed to operate at 100 kHz for the detection of flaws in stainless steel tubes used in nuclear reactor steam generators. The coil on the bottom is used for the detection of cracks and flaws in air conditioning tubing (8 mm) and operates at 200 kHz. The output is the difference between the outputs of the two coils.

FIGURE 5.30 ■ Inductive proximity sensors.

FIGURE 5.31 ■ Eddy current sensors. Top: 100 kHz absolute sensor embedded in a dielectric medium. Bottom: 400 kHz differential sensor embedded in ferrite.

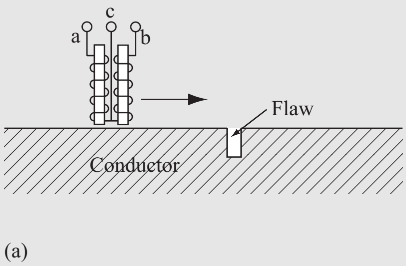
FIGURE 5.32 ■ Eddy current sensors (differential) for the detection of flaws in tubing. Top: 19 mm, 100 kHz sensor for stainless steel tubing in nuclear power plant steam generators. Bottom: 8 mm, 200 kHz sensor for air conditioning tubing.

EXAMPLE 5.6 Eddy current testing for flaws

The idea of position sensing can be used for nondestructive testing of materials to help in the detection of cracks, holes, and subsurface anomalies in conducting media. Consider the configuration in **Figure 5.33a**, where a thick aluminum conductor has a 2.4 mm hole drilled in it to some depth, representing a flaw. Two small coils, each 1 mm in diameter and separated 3 mm apart, are placed against the aluminum surface and slid to the right in small increments (the bottom probe in **Figure 5.31** was used for these measurements). The inductance of each coil is measured and the difference between the two inductances is used as an indication. This measurement is differential and is particularly useful when the environment is noisy.

Figure 5.33b shows a plot of the inductance versus position of the center of the probe, starting about 18 mm from the center of the hole and moving 18 mm past the hole. Since the two probes are identical, their inductance is identical, and as long as they see identical conditions, the output is zero. Far from the hole, or when the probe is centered over the hole, the output is zero. Anywhere else, one coil will have a higher or a lower inductance than the other and hence a variable output. Because the difference is taken between the leading coil and the trailing coil, and since the inductance in the vicinity of the flaw is lower, the curve first dips (negative difference) then rises to a positive difference.

In actual tests, the coils are fed with a constant AC and the potential across the two coils is measured. The potential is complex, but will vary in the same manner as the inductance.



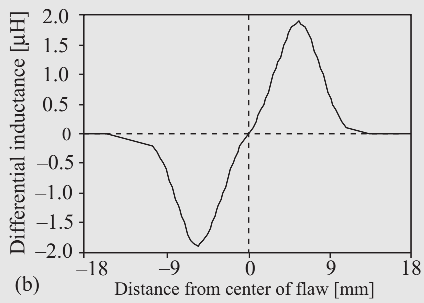


FIGURE 5.33 ■ Differential probe, eddy current nondestructive testing. (a) The probe and geometry. (b) The output given as the difference in inductance between the leading and trailing coils.

5.4.1.3 *Position and Displacement Sensing: Variable Inductance Sensors*

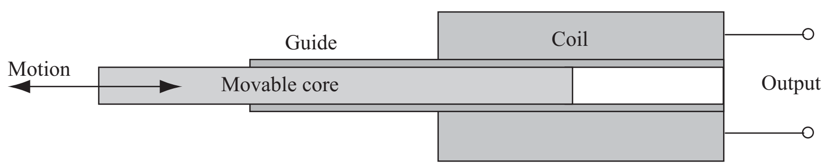
Position and displacement are usually understood as measuring the exact distance from a point or the travel of a point relative to another. This requires accurate measurements and possibly linear transfer functions of the sensors involved. One approach to this task is through the use of variable inductance sensors, sometimes called variable reluctance sensors. Magnetic reluctance is the equivalent magnetic term to electric resistance and is defined as (see also **Equation (5.32)**):

$$\mathcal{R} = \frac{L}{\mu S} \quad [1/\mathrm{H}]. \tag{5.36}$$

The reluctance is smaller the shorter the magnetic path, the larger its cross-sectional area, and the larger its permeability. Reluctance is then related to inductance through permeability, and reducing reluctance increases inductance and vice versa. Typically the reluctance of a coil can be changed by adding a gap in the magnetic path and changing the effective length of this gap.

Thus one of the simplest methods of changing the inductance of a coil is to provide it with a movable core as shown in **Figure 5.34**. In this sensor, the further the movable core moves in, the smaller the reluctance of the magnetic path and the larger

FIGURE 5.34 ■ Inductive sensor with a movable core.



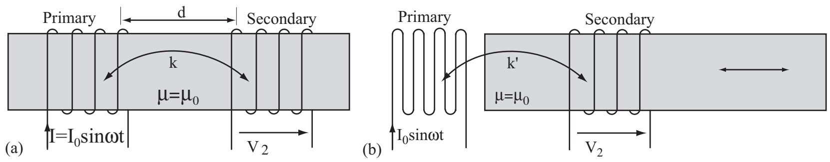


FIGURE 5.35 ■ Principle of the LVDT. (a) Changing the distance between the coils. (b) Moving the core between two fixed coils. In either case, the coupling coefficient k changes.

the inductance. If the core is made of a ferromagnetic material, the inductance increases, whereas a conducting material reduces the inductance (see explanation for eddy current proximity sensors). This type of sensor is called a ***linear variable inductance sensor***. Linear here means that the motion is linear, not necessarily that the transfer function is linear. By sensing the inductance, a measure of the position of the core is available. The same configuration may be used to measure force, pressure, or anything else that can produce linear displacement.

A better displacement sensor is a sensor based on the idea of the transformer. This is based on one of two related principles: either the distance between two coils of a transformer is varied (the coupling between the coils changes) or the coupling coefficient between the two coils is varied by physically moving the core while the two coils are fixed. Both principles are shown in **Figure 5.35**. A variation of the second of these is the **linear variable differential transformer** (LVDT), which will be discussed shortly. To understand the principles, first consider **Figure 5.35a**. Assuming a constant AC voltage V_{ref} is connected across the primary coil, the induced voltage in the secondary coil is the output voltage:

$$V_{out} = k\frac{N_2}{N_1}V_{ref} \quad [\mathrm{V}]. \tag{5.37}$$

k is a coupling coefficient that depends on the distance between the coils as well as the medium between them and any other material that may be present in the vicinity, such as a shield, enclosure, etc. Given a calibration curve, the output voltage, which can be measured directly, is a measure of the distance between the coils. In **Figure 5.35b**, the same relation holds except that now the moving core changes the coupling between the coils, thus changing the output voltage across the second coil.

Another type of position sensor may be made by connecting the two coils in **Figure 5.35a** in series and measuring the inductance of the two coils. The latter is $L_{11} + L_{22} + 2L_{12}$, where L_{12} is the mutual inductance between them, given as $L_{12} = k\sqrt{L_{11}L_{22}}$. The coupling coefficient k depends on the distance between the two coils, and by measuring the total inductance, one gets a measure of the position of one coil with respect to the second. This arrangement is called a ***coil-displacement sensor***.

The LVDT sensor is shown in **Figure 5.36**. This is similar to the variable inductance sensor, but now there are two coils in the output circuit whose voltages subtract. When the core is symmetrical about the coils (**Figure 5.36a**), the output of the sensor is zero. If the core moves to the left, the voltage on coil 2 decreases (**Figure 5.36b**) while the voltage on coil 1 remains the same, since only the coupling between the reference (primary) coil and coil 2 changes. The total voltage now increases and its phase is, say, positive. When the core moves to the right, the opposite happens, but the phase is

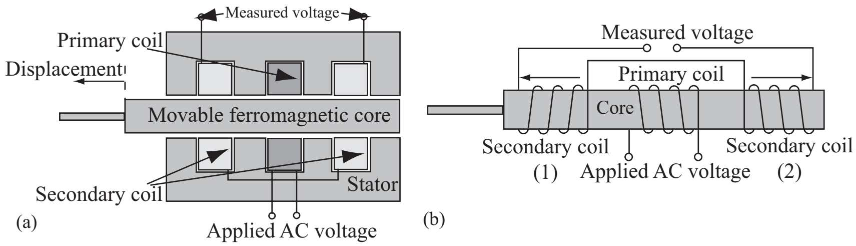


FIGURE 5.36 ■ The moving core LVDT. (a) Construction. (b) Principle.

opposite (negative in this case). This change in phase is used to detect the direction of motion, whereas the output voltage is a measure of distance the core travels from the zero output position. These devices are very sensitive and useful, and in a relatively small range of motion, the output is linear. The reference coil is driven with a stable sinusoidal source at a constant frequency and the core is ferromagnetic. The whole sensor is enclosed and shielded so that no field extends outside it and hence cannot be influenced by outside fields. The core slides in and out, and that motion is often used for accurate measurements of displacement for applications in industrial control and machine tools.

LVDT sensors are extremely rugged and come in various dimensions to suit many needs (some as small as 10 mm long). In most practical applications, the voltage output is measured (amplification is usually not needed), whereas the phase is detected with a zero-crossing phase detector (a comparator; see **Chapter 11**). The frequency of the source must be high enough with respect to the frequency of motion of the core (a figure of 10 times higher is common) to avoid errors in the output voltage due to slow response of the LVDT. The operation of LVDTs can be from AC sources or DC sources (with an internal oscillator providing the sinusoidal voltage). Typical working voltages are up to about 25 V, whereas output is usually below 5 V. Resolution can be very high, with a linear range about 10%–20% of the length of the coil assembly. Although the stated function of LVDTs is position sensing, anything that can be related to position can also be sensed. One can use an LVDT to sense fluid level, pressure, acceleration, and many other functions.

A variation of the LVDT is the ***rotary variable differential transformer*** (RVDT), intended for angular displacement and rotary position sensing. In all respects it is identical in operation to the LVDT device, but the rotary motion imposes certain restrictions on its construction. The RVDT is shown schematically in **Figure 5.37** and includes a ferromagnetic core that couples with the secondary coils based on angular position. The moving core is shaped to obtain linear output over the useful range of the

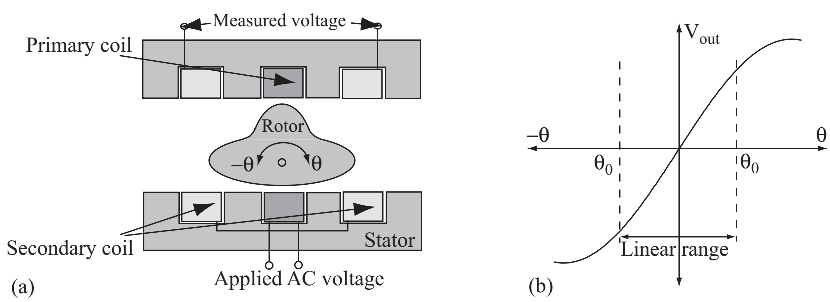


FIGURE 5.37 ■ Schematic view of an RVDT. (a) Construction. (b) Output.

sensor. Other arrangements are possible as well. The span is given in angles and can be up to $\pm 40°$. Beyond that the output becomes nonlinear and less useful.

5.4.2 Hall Effect Sensors

The Hall effect was discovered in 1879 by Edward H. Hall. The effect exists in all conducting materials, but is particularly pronounced and useful in semiconductors. To understand the principle, consider a block of conducting medium through which a current of electrons is flowing caused by an external source, as shown in **Figure 5.38**. A magnetic flux density B is established across the conductor, making an angle θ with the direction of the current (in **Figure 5.38**, $\theta = 90°$). The electrons flow at a velocity v, and according to **Equation (5.22)**, a force perpendicular to both the current and field is established on the flowing electrons. Since the force is related to the electric field intensity as $F = qE$, we can write the electric field intensity in the conductor as

$$E_H = \frac{F}{q} = vB \sin\theta \quad [\mathrm{V/m}]. \tag{5.38}$$

The index H indicates this is the Hall electric field and it is perpendicular to the direction of current flow. To rewrite this in terms of current flowing in the element, we note that the current density may be written as $J = nqv$ [A/m^2], where nq is the charge density (n is number of electrons per cubic meter and q is the charge of the electron). v is the average velocity of electrons. Therefore the Hall electric field intensity is

$$E_H = \frac{nqvB \sin\theta}{nq} = \frac{JB \sin\theta}{nq} \quad [\mathrm{V/m}]. \tag{5.39}$$

The current density is the current I divided by the cross-sectional area perpendicular to the flow of current or $J = I/Ld$:

$$E_H = \frac{IB \sin\theta}{nqLd} \quad [\mathrm{V/m}]. \tag{5.40}$$

The force pulls the electrons toward the front surface of the conductor and therefore a voltage develops between the back (positive) and front (negative) surfaces. The potential

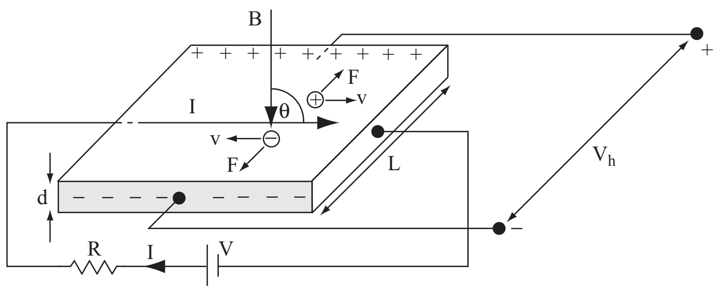


FIGURE 5.38 ■ The Hall element. The two opposite electrodes on the far and near surfaces measure the Hall voltage while a current flows across horizontally. The component of the magnetic flux density perpendicular to the element (shown here as B) is the sensed quantity.

difference is the integral of the electric field intensity E along the path of length L. Because Hall elements are typically small, we may assume E is constant along L and write

$$V_H = EL = \frac{IB \sin \theta}{qnd} \quad [\text{V}]. \tag{5.41}$$

This voltage is the ***Hall voltage***. In particular, in measurements, the angle θ is typically 90° and the Hall voltage is given by

$$V_H = \frac{IB}{qnd} \quad [\text{V}], \tag{5.42}$$

where d is the thickness of the hall plate, n is the carrier density [charges/m^3], and q is the charge of the electron [C].

It should be noted that if the current changes direction or the magnetic field changes direction, the polarity of the Hall voltage flips. Thus the Hall effect sensor is polarity dependent, a property that may be used to good advantage to measure the direction of a field or the direction of motion if the sensor is properly set up.

The term $1/qn$ [m^3/C] (or [m^3/A·s]) is material dependent and is called the **Hall coefficient** (K_H):

$$K_H = \frac{1}{qn} \quad [\text{m}^3/\text{A·s}]. \tag{5.43}$$

Strictly speaking, K_H in conductors is negative since q is the charge of the electron. The **Hall voltage** is usually represented as

$$V_{\text{out}} = K_H \frac{IB}{d} \quad [\text{V}]. \tag{5.44}$$

The relations above apply to all conductors. In semiconductors the Hall coefficient depends on both hole and electron mobilities and concentration as follows:

$$K_H = \frac{p\mu_h^2 - n\mu_e^2}{q(p\mu_h + n\mu_e)^2} \quad [\text{m}^3/\text{A·s}], \tag{5.45}$$

where p and n are the hole and electron densities, respectively, μ_h and μ_e are the hole and electron mobilities, respectively, and q is the charge of the electron. The net effect of this dependency is a large coefficient, so much so that all practical Hall sensors are based on semiconductors. **Equations (5.44)** and **(5.45)** may in fact be used to measure properties of materials such as charge densities and mobilities based on the Hall voltage. It should also be noted from **Equation (5.45)** that the hole and electron densities will affect the Hall coefficient. Large doping with n-type dopants will produce a negative coefficient, whereas large p-type doping will make the coefficient positive. There is a certain doping level at which the coefficient is zero, as can be calculated directly from **Equation (5.45)**.

The Hall coefficient can also be related to the conductivity of the medium. Since conductivity is related to the mobility of charges, for conductivity in conductors we have

$$\sigma = nq\mu_e \quad [\text{S/m}]. \tag{5.46}$$

In semiconductors the conductivity depends on the mobility of both electrons and holes:

$$\sigma = qn\mu_e + qp\mu_h \quad [\mathrm{S/m}]. \tag{5.47}$$

Therefore the Hall coefficient in conductors can be written as

$$K_H = \frac{\mu_e}{\sigma} \quad [\mathrm{m^3/A{\cdot}s}]. \tag{5.48}$$

In semiconductors we have

$$K_H = \frac{q\left(p\mu_h^2 - n\mu_e^2\right)}{\sigma^2} \quad [\mathrm{m^3/A{\cdot}s}]. \tag{5.49}$$

In principle, the lower the conductivity, the higher the Hall coefficient. However, this is only true to a point. As the conductivity decreases, the resistance of the device increases and the current in the device decreases, reducing the Hall voltage (see **Equation (5.44)**).

It should also be recalled here that in doped semiconductors, the product of the electron and hole concentrations is related to the intrinsic concentration through the mass action law:

$$np = n_i^2. \tag{5.50}$$

An intrinsic material is that in which $n_i = n = p$.

These relations clearly indicate that the Hall effect can be used to measure conductivity or, alternatively, that any quantity that affects the conductivity of the medium also affects the Hall voltage. For example, a semiconducting Hall element exposed to light will read in error due to the change in conductivity of the semiconductor due to the photoconducting effect (see **Section 4.4.2.2**).

Hall coefficients vary from material to material and are particularly large in semiconductors. For example, in silicon it is on the order of -0.02 m^3/A·s, but it depends on doping and temperature, among other parameters. The most important aspect of this sensor is that it is linear with respect to the field for a given current and dimensions. However, the Hall coefficient is temperature dependent, and this must be compensated for if accurate sensing is needed. Because the Hall voltage in most materials is rather small—on the order of 50 mV/T—and considering the fact that most fields sensed are smaller than 1 T, the Hall voltage must in almost all cases be amplified. As an example, the earth's magnetic field is only about 50 μT, so the output of a Hall sensor in the terrestrial magnetic field is on the order of 25 μV. Nevertheless, these are easily measurable quantities and Hall sensors are among the most commonly used sensors for magnetic fields because they are simple, linear, can be integrated within semiconductor devices, and are inexpensive. They are available in various forms, sizes, sensitivities, and in arrays. The errors involved in measurement are mostly due to temperature variations, but the size of the Hall plate, if large, also introduces averaging errors due to its integration effect. Some of these effects can be compensated for by appropriate circuitry or compensating sensors. In terms of fabrication, a typical sensor is a thin rectangular wafer made of *p*- or *n*-doped semiconductor (InAs and InSb are the most commonly used materials because of their larger carrier densities, and hence larger Hall coefficients, but silicon may also be used with reduced sensitivity). The sensor is usually identified by two

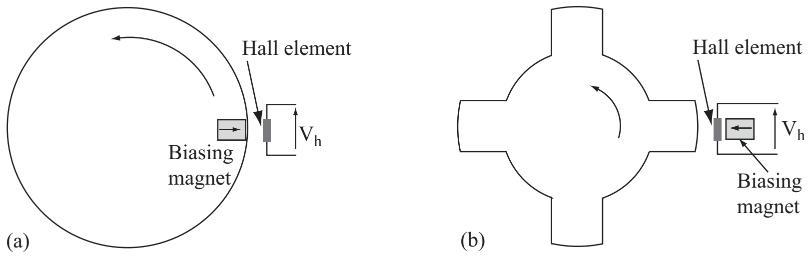


FIGURE 5.39 ■ (a) Sensing rotation of a shaft. The small magnet (arrow shows the direction of the field) induces a voltage pulse V_h every time it passes past the Hall element. (b) The Hall element senses the position of the rotating poles in a 4-cylinder engine to fire the appropriate cylinder in sequence and at the correct time.

transverse resistances: the control resistance through which the control current flows and the output resistance across which the Hall voltage develops.

In practical applications, the current is usually kept constant so that the output voltage is directly proportional to the field. The sensor may be used to measure the flux density (provided proper compensation can be incorporated) or may simply be used as a detector or operate a switch. The latter is very common in the sensing of rotation, which, in itself, may be used to measure a variety of effects (shaft position, frequency of rotation [rpm], differential position, torque, etc.). An example is shown in **Figure 5.39a**, where the rotation of a shaft is sensed. An emf is induced in the Hall element every time the small magnet passes by indicating the rotation of the shaft. Many variations of this basic configuration can be envisioned, including, for example, the measurement of angular displacement. Hall elements are integral to many electrical motors in disk drives, CD-ROM drives, and many other applications in which the rotation is sensed and controlled through these sensors.

Hall elements are also fabricated in pairs, separated a small distance apart, for the expressed purpose of sensing the gradient in the field rather than the field itself. This is particularly useful in position and presence sensing in which the edge of a ferromagnetic medium, such as gears in transmissions or in electronic ignition systems, is sensed. Some sensors come with their own biasing magnet to generate the magnetic field and may have either analog or digital output with onboard electronics. In these devices, changes in the magnetic field due to the presence of a ferromagnetic material are sensed. An example of this type of sensor used for position sensing is shown in **Figure 5.39b**. This particular configuration is common in electronic ignition systems (a 4-cylinder application is shown), where a pulse is produced by the Hall element every time one of the metal poles passes by. The material of the poles must be ferromagnetic (iron). This configuration, in many variants and for both linear and angular applications, is one of the simplest methods of sensing position.

The Hall sensor can be used for other applications. One example is the direct sensing of electric power. In an application of this type (**Figure 5.40**), the power, which is the product of current and voltage, is measured as follows: The voltage is connected to a coil that generates a magnetic field across the Hall element. The current, which now can be variable, is the control current in the sensor. The Hall voltage is proportional to power and, if properly calibrated, will measure power directly.

Hall element sensors are usually considered to be DC devices. Nevertheless, they can be easily used to sense alternating fields at relatively low frequencies. The specification sheet for Hall elements gives their response and the maximum useful frequency.

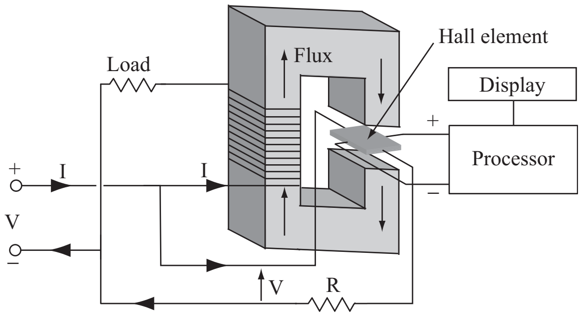


FIGURE 5.40 ■ Direct sensing of power with a single Hall element.

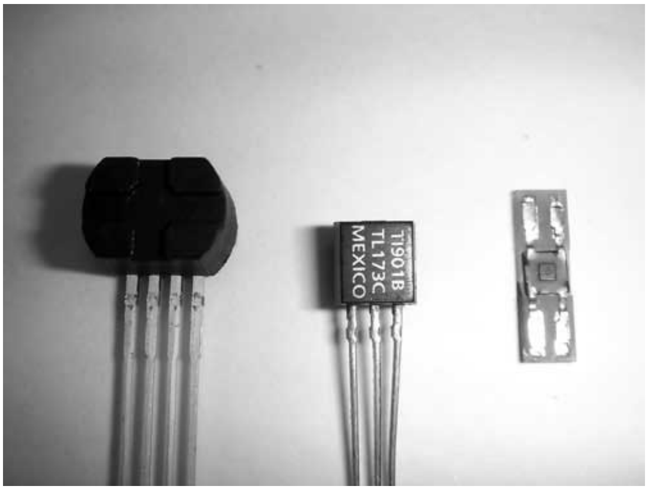


FIGURE 5.41 ■ Various Hall elements. Left: A dual Hall element with biasing magnet and digital output. Middle: An analog Hall sensor. Right: A Hall element chip, mounted on a small fiberglass piece. The chip is 1 mm × 1 mm.

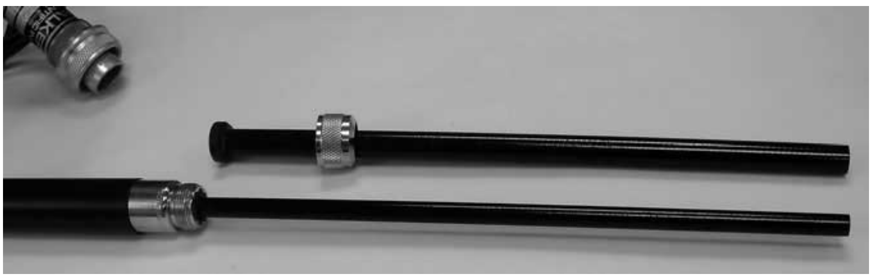


FIGURE 5.42 ■ A Hall element probe. The tip of the probe (bottom, right) contains three Hall element sensors at right angles to measure the three components of the magnetic field.

Figures 5.41 and **5.42** show a number of Hall elements/sensors, including a probe for three-axis measurement of fields down to 10 μT. The latter has three sensors, placed at right angles, and fed to three separate amplifiers.

Finally, it should be reemphasized that the only quantity that Hall sensors measure directly is the magnetic flux density, but they can be made to sense a whole range of quantities by judicious use of the sensors in conjunction with mechanical and electrical arrangements, the examples in **Figures 5.39** and **5.40** being representative examples.

EXAMPLE 5.7 Measurement of magnetic flux density and magnetic flux using a Hall element

The Hall element's primary function is sensing of magnetic fields, but it can also sense any quantity related to the magnetic field. The geometry in **Figure 5.43** uses a silicon element with a Hall coefficient of -10^{-2} m^3/A·s. The dimensions of the Hall element are $a = 2$ mm, $b = 2$ mm, and its thickness is $c = 0.1$ mm. Calculate the response of the Hall element:

a. For magnetic flux densities from 0 to 2 T. This is the range normally found in electric machines. What is the minimum field measurable if a digital voltmeter with a resolution of 2 mV is used to measure the Hall voltage?
b. For flux from 0 to 10 μWb.

Solution: The magnetic flux density is measured directly as indicated in **Equation (5.44)**. The flux is not measured directly, but rather, since $\Phi = B \times S$, we still measure the flux density B and translate the measurement into flux. Thus

a. The relation between flux density and the Hall voltage is

$$V_{\text{out}} = K_H \frac{I_H B}{d} = 0.01 \times \frac{5 \times 10^{-3}}{0.1 \times 10^{-3}} B = 0.5B \text{ V}.$$

This is a linear transfer function varying from 0 to 1 V for a flux density varying from 0 to 2 T. The sensitivity of the device is clearly 0.5 V/T. A 1 mV corresponds to 1 T/500 = 0.002 T. Thus a 2 mV voltmeter will measure a minimum flux density of 4 mT. This is not particularly sensitive (i.e., 4 mT = 4000 μT is much higher than the terrestrial magnetic flux density of 60 μT), but it is useful for higher fields.

b. To measure flux, we recall that flux is flux density integrated over area. Since the area of the sensor is small, $S = 4 \times 10^{-6}$ m^2, we may assume the flux density to be constant over the area and simply multiply the flux density by area:

$$\Phi = BS \quad [\text{Wb}].$$

But since we sense the flux density, the Hall voltage sensed will be

$$V_{out} = K_H \frac{I_H B}{d} = K_H \frac{I_H BS}{Sd} = K_H \frac{I_H \Phi}{Sd} = 0.01 \times \frac{5 \times 10^{-3}}{0.1 \times 10^{-3} \times 4 \times 10^{-6}} \Phi = 1.25 \times 10^5 \Phi \text{ V}.$$

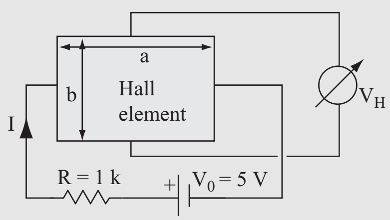


FIGURE 5.43 ■ Biasing a Hall element. The magnetic flux density is perpendicular to the Hall element plate.

The sensitivity is 1.25×10^5 V/Wb. For the range between 0 and 10 μWb, the output voltage will vary from 0 to $1.25 \times 10^5 \times 10 \times 10^{-6} = 1.25$ V. The minimum measurable flux using the same voltmeter is $10/625 = 0.016$ μWb.

EXAMPLE 5.8 Rotation speed of an engine

The rotation speed of an engine needs to be sensed for the purpose of speed regulation. To do so, a Hall element is used as a sensor using the configuration in **Figure 5.44a**. Two symmetric bumps or protrusions are added to the shaft (two bumps are needed to keep its mass balanced). The gap between the Hall element and the shaft varies from 1 mm when one of the bumps is aligned with the Hall element to 2 mm when it is not. Calculate the minimum and maximum reading of the Hall element if it is biased using the circuit in **Figure 5.44b**, which has a Hall coefficient of 0.01 m^3/A·s and is 0.1 mm thick. Assume the permittivity of the shaft and the iron ring are very high and the permeability of the Hall element is the same as air, equal to μ_0. The coil contains 200 turns and is supplied with 0.1 A to produce a magnetic flux density in the gap.

Solution: Because the permeability of the iron ring is large, we can neglect its reluctance, meaning that the flux in the gap only depends on the gap length. From **Equation (5.36)** we write for the gap:

$$\mathcal{R}_g = \frac{l_g}{\mu_0 S},$$

where S is the cross-sectional area of the gap. We do not have that quantity, but it is not necessary since we calculate the flux and then divide by the area S to find the magnetic flux density. The flux in the gap is

$$\Phi_g = \frac{NI}{\mathcal{R}_g} = \frac{NI\mu_0 S}{l_g}.$$

The flux density thus becomes

$$B_g = \frac{\Phi_g}{S} = \frac{NI\mu_0}{l_g}.$$

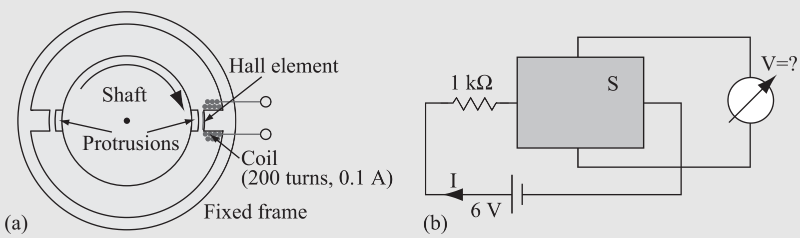


FIGURE 5.44 ■ An engine rotation speed sensor. (a) Geometry including the magnetizing coil and Hall element. (b) Electrical connections for the Hall element.

Now from **Equation (5.44)** we write

$$V_{out} = K_H \frac{I_H B}{d} = K_H \frac{I_H}{d} \frac{NI\mu_0}{l_g} \quad [\mathrm{V}],$$

where I_H is the bias current in the Hall element (6 mA in this case) and I is the current in the coil (0.1 A). For the 1 mm gap we get (there are two identical gaps, one on each side):

$$V_{\max} = 0.01 \times \frac{6 \times 10^{-3}}{0.1 \times 10^{-3}} \times \frac{200 \times 0.1 \times 4\pi \times 10^{-7}}{2 \times 1 \times 10^{-3}} = 0.0075 \text{ V}.$$

For the 2 mm gap we get

$$V_{\min} = 0.01 \times \frac{6 \times 10^{-3}}{0.1 \times 10^{-3}} \times \frac{200 \times 0.1 \times 4\pi \times 10^{-7}}{2 \times 2 \times 10^{-3}} = 0.00375 \text{ V}.$$

The output voltage varies from 3.75 mV when the Hall element is not in the gap to 7.5 mV when it is. The output is a signal approximating a sinusoidal signal varying from a peak of 7.5 mV to a valley of 3.75 mV at a frequency twice the rate of rotation of the shaft. For example, an engine shaft rotating at 4000 rpm produces a frequency of $(4000/60) \times 2 = 133.33$ Hz. Measurement of this frequency (possibly after amplifying the signal and digitization) and proper calibration of the reading produces the necessary data.

5.5 | MAGNETOHYDRODYNAMIC (MHD) SENSORS AND ACTUATORS

The force relations in **Equations (5.21)** and **(5.25)** may be used in a number of ways for both sensing and actuation in addition to those discussed above. Because the magnetic forces involved act on charges, and therefore on currents, these are responsible for most methods of magnetic actuation and, as we have seen from the Hall effect, for sensing as well. One particularly interesting aspect of the magnetic force is the ability to create forces in moving media such as plasmas, charged gases, and liquids, and also in solid conductors. The phenomenon has been dubbed magnetohydrodynamics because of its use in moving charged gases, fluids, and molten metals, but the principle is fundamentally the same as used in electric motors and generators. The principles involved are demonstrated in **Figure 5.45a**, which shows an MHD generator (i.e., a sensor), and in **Figure 5.45b**, which shows an MHD pump or actuator. Both involve a channel containing a conducting medium. The medium may be any medium, provided it contains

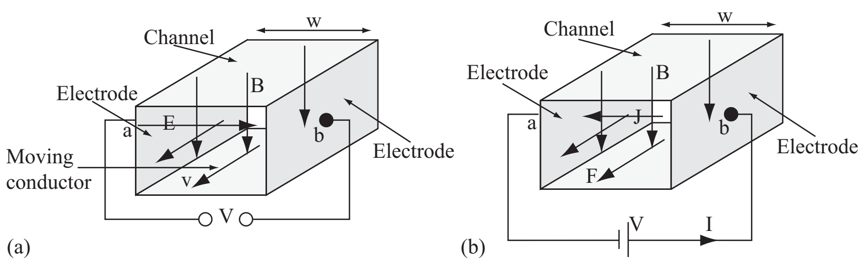


FIGURE 5.45 ■ (a) An MHD generator (sensor). (b) An MHD pump or actuator.

charges on which the magnetic field acts. The charges may be free charges in a conductor or a charged plasma.

5.5.1 MHD Generator or Sensor

The medium in the channel in **Figure 5.45a** moves at a velocity v, as shown. The magnetic field acts on the moving charges based on **Equation (5.21)**:

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} = q\mathbf{E} \quad [\mathrm{N}], \tag{5.51}$$

where the fact that the force on a charge may be written in terms of the electric field intensity that produces that force has been added. This means that the moving charged medium produces an electric field intensity, as shown in the figure by the horizontal arrow. Consequently, a potential difference is generated across the two electrodes. The latter is

$$V_{ba} = -\int_a^b \mathbf{E} \times d\mathbf{l} = -\int_a^b \mathbf{v} \times \mathbf{B} \times d\mathbf{l} = vBw \quad [\mathrm{V}]. \tag{5.52}$$

This is the basic principle of MHD generation and that of sensing: the velocity of the charged medium can be sensed by measuring the MHD voltage generated across the electrodes. However, for the potential to develop, there must be charges that can be separated, or on a macroscopic level, we may say that the medium must have a nonzero conductivity.

5.5.2 MHD Pump or Actuator

The generation process can be reversed by passing a current through the channel, perpendicular to the magnetic field, as shown in **Figure 5.45b**. The current density in the channel creates a force on the conducting medium, forcing it out of the channel based on **Equation (5.25)**:

$$\mathbf{F} = \int_v \mathbf{J} \times \mathbf{B} dv \quad [\mathrm{N}]. \tag{5.53}$$

This force may be used to pump any conducting liquid (such as molten metals or seawater), a conducting gas (plasma), or a solid conductor.

The main appeal of MHD actuators is that the system is trivially simple and requires no moving elements. It can, under the right conditions, generate enormous forces (and therefore accelerations). On the other hand, apart from some applications in pumping molten metals, it is an inefficient method of both generation and actuation. Nevertheless, it has found applications in sensing, where the efficiency of the generation process is not important, and in a number of actuators, including pumping, particle acceleration, and acceleration of solid objects in so-called rail guns, devices proposed for military and space applications.

An MHD flow sensor is shown in **Figure 5.46**. The two coils, one on top and one on the bottom, generate a magnetic flux density between the top and bottom surfaces (this can be done equally well with two permanent magnets). The fluid, such as water, must have free ions (primarily Na^+ and Cl^-) for the system to work. Fortunately most fluids, including water, have sufficient dissolved salts to provide ions for MHD sensing. The output is directly proportional to the velocity of the fluid and the magnetic flux density

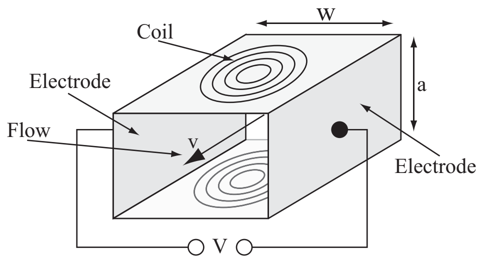


FIGURE 5.46 ■ An MHD flow rate sensor.

applied. Assuming that the flux density is constant throughout the width of the channel, the voltage measured is proportional to the fluid velocity:

$$V = Bwv \quad [\mathrm{V}]. \tag{5.54}$$

The sensor can also sense flow rate (volume/second) as

$$Q = wav = \frac{aV}{B} \quad [\mathrm{m^3/s}]. \tag{5.55}$$

The same configuration can be used, for example, to sense the velocity of a boat relative to water, especially in seawater.

The basic actuation method in **Figure 5.45b** can be used in a number of ways. **Figure 5.47** shows a pump for molten metals (aluminum, magnesium, sodium, etc.). The molten metal flows in a conduit equipped with two side electrodes, as shown. The force on the molten conductor is given by **Equation (5.53)**. In this case, the magnetic flux density, $\mathbf{B}_0$, is generated by a current in the coil, I_0. The volume of integration is that part of the volume of the pump in which the current I, produced by a separate source, and magnetic flux density interact, or approximately abd. Therefore, assuming that the current is uniform in this section and the power supply generates a current I, the current density is approximately I/bd, whereas the coil and armature generate a flux density B_0 and the force can be estimated as

$$F = B_0 \frac{I}{bd} abd = B_0 I a \quad [\mathrm{N}]. \tag{5.56}$$

The reason this force is only an approximation is that the flux density and the current density are not necessarily constant within the volume shown, but nevertheless, the result is a good approximation, especially in highly conducting media such as molten metals. The pump shown here has a significant advantage over other methods in that there is no mechanical interaction with the molten metal. However, it should also be

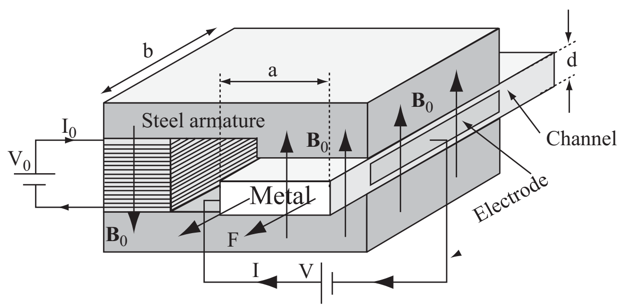


FIGURE 5.47 ■ An MHD pump for molten metals. The metal is contained within an enclosed channel.

remembered that the device, especially the coil, must be properly cooled and that it consumes a significant amount of power.

It should be noted that the force in **Equation (5.56)** is the same force that would apply on a current-carrying wire of length a, carrying a current I, in a magnetic field B_0, as was found previously in **Equation (5.26)**. This means that the same forces act in electric machines, but the designation MHD helps in separating applications that look very different.

EXAMPLE 5.9 Electromagnetic propulsion in seawater

A simple method of propulsion of vessels in seawater has been proposed based on MHD pumping. The dimensions of an actuator for this purpose are shown in **Figure 5.48**. The magnetic field is generated using a permanent magnet, producing a constant magnetic flux density $B = 0.6$ T. The conductivity of seawater is 4 S/m.

a. With the dimensions shown, calculate the power needed to generate 10 tons of force (10^5 N) to propel the vessel in seawater.
b. As a comparison, suppose that the same device is used to pump molten sodium (at 98°C) to cool a nuclear reactor. What is the power needed to produce the same force? The conductivity of molten sodium is 2.4×10^7 S/m.

Solution:
a. The force is given in **Equation (5.56)**.

$$F = B_0 Ia \quad [\mathrm{N}].$$

Since B_0 is known, we need to calculate the necessary current. For 10 tons or 100,000 N we get

$$I = \frac{F}{Ba} = \frac{100000}{0.5 \times 0.8} = 250,000 \;\; \mathrm{A}.$$

To calculate power we need to calculate the resistance of the channel. The latter is

$$R = \frac{a}{\sigma bc} = \frac{0.5}{4 \times 0.25 \times 1} = 1 \; \Omega.$$

The power required is 62.5 GW.

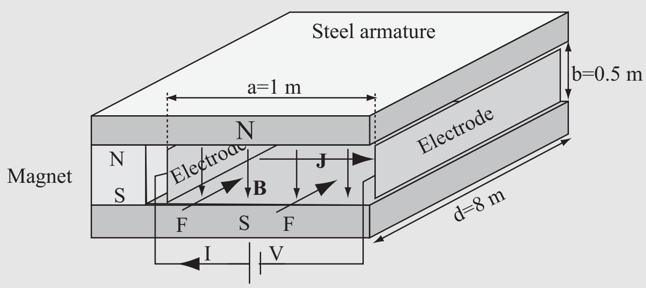


FIGURE 5.48 ■ MHD propulsion.

Clearly this is not a practical device. A 62.5 GW power source that generates 10 tons of thrust is not a viable system. The voltage source would have to supply 250 kV DC, something that is not trivial to generate even if one neglects the power needs. However, the method is valid under the right conditions, as part (b) shows.

b. What changes here is the resistance of the channel:

$$R = \frac{a}{\sigma bc} = \frac{0.5}{2.4 \times 10^7 \times 0.25 \times 1} = 8.33 \times 10^{-8}\ \Omega.$$

Therefore the power required is

$$P = I^2 R = \frac{a}{\sigma bc} = (250,000)^2 \times 8.33 \times 10^{-8} = 5.208 \times 10^3\ \mathrm{W},$$

or just over 5.2 kW. The voltage of the source is only 20 mV. Low-voltage, high-current sources are easier to build than high-voltage, high-current sources, although the case here is somewhat extreme. Nevertheless, one can use a transformer to reduce an AC voltage to a required level and then rectify it or one can produce this low-voltage, high-current source with special (homopolar) generators.

5.6 | MAGNETORESISTANCE AND MAGNETORESISTIVE SENSORS

Magnetoresistance is the effect the magnetic field has on the electrical resistance of a conductor or semiconductor. There are two mechanisms through which the resistance of a medium changes in the presence of a magnetic field. The first is due to the fact that electrons are attracted or repelled by magnetic fields, as was discussed above in conjunction with the Hall effect. The second mechanism exists in certain materials in which the direction of internal magnetization due to flow of a current changes with the application of an external magnetic field. This manifests itself in changes in the electrical resistance of the medium.

Magnetoresistors based on the first of these mechanisms are very similar to the Hall elements in that the same basic structure of **Figure 5.38** is used, but without the Hall voltage electrodes. The effect exists in all materials but is very pronounced in semiconductors. This is shown in **Figure 5.49a**. The electrons are affected by the magnetic field, as in the Hall element, and because of the magnetic force on them, they will flow in an arc, as shown in **Figure 5.49b**. The larger the magnetic flux density, the larger the force exerted on the electrons and the larger the radius of the arc. This effectively forces the electrons to take a longer path, which means that the resistance to their flow increases (exactly the same as if the effective length of the plate is larger). Thus a relationship between the magnetic field and current is established. This relation is

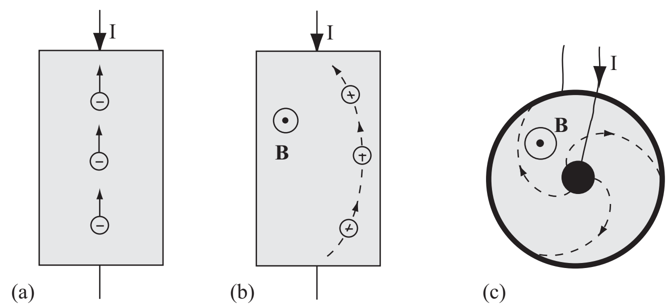


FIGURE 5.49 ■ Magnetoresistance in a semiconductor. (a) No magnetic field. (b) A magnetic field alters the flow path of the carriers. (c) The Corbino disk magnetoresistor.

proportional to B^2 for most configurations and is dependent on carrier mobility in the material used (usually a semiconductor). However, the exact relationship is rather complicated and depends on the geometry of the device. Therefore we will simply assume here that the following holds:

$$\frac{\Delta R}{R_0} = kB^2, \tag{5.57}$$

where k may be viewed as a calibration function. A particularly useful configuration for magnetoresistor is shown in **Figure 5.49c**. This is called the Corbino disk and has one electrode at the center of the disk and a second electrode on the perimeter. By so arranging the electrodes, the sensitivity of the device increases because of the long spiral paths electrons take in flowing from one electrode to the other.

Magnetoresistors are used in a manner similar to Hall elements, but their use is simpler since one does not need to establish a control current. Rather, the measurement of resistance is all that is necessary. The device is a two-terminal device built from the same types of materials as Hall elements (InAs and InSb in most cases). Magnetoresistors are often used where Hall elements cannot be used. One important application is in magnetoresistive read heads where the magnetic field corresponding to recorded data is sensed.

The second type of magnetoresistive sensor, which is even more sensitive than the basic element discussed above, is based on the property of some materials to change their resistance in the presence of a magnetic field when a current flows through them. Unlike the sensors discussed above, these are metals with highly anisotropic properties and the effect is due to the change in their magnetization direction due to the presence of a magnetic field. The effect is called **anisotropic magnetoresistance** (AMR), discovered in 1857 by William Thomson (Lord Kelvin).

One of the structures most commonly used in commercial magnetoresistive sensors is the following: A magnetoresistive material is exposed to the magnetic field to be sensed. A current passes through the magnetoresistive material at the same time. This is shown schematically in **Figure 5.50**. The magnetic field is applied perpendicular to the current. The sample has an internal magnetization vector parallel to the flow of current. When the magnetic field is applied, the internal magnetization changes direction by an angle α, as shown in **Figure 5.50**. The resistance of the sample becomes

$$R = R_0 + \Delta R_0 \cos^2\alpha, \tag{5.58}$$

where R_0 is the resistance without application of the magnetic field and ΔR_0 is the change in resistance expected from the material used. Both of these are properties of the material and the construction (for R_0). The angle α is proportional to the applied field and is material dependent. Some anisotropic magnetoresistive materials and their properties are shown in **Table 5.4**.

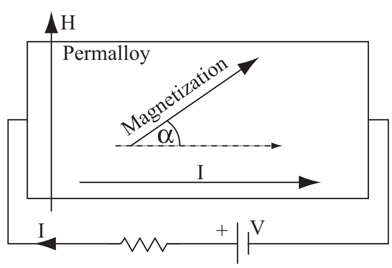


FIGURE 5.50 ■ Principle of operation of an AMR sensor.

TABLE 5.4 ■ Some AMR materials and some of their properties

Material	Resistivity $\rho = (1/\sigma) \times 10^{-8}$ Ωm	$\Delta\rho/\rho$ %
Fe19Ni81	22	2.2
Fe14Ni86	15	3
Ni50Co50	24	2.2
Ni70Co30	26	3.2
Co72Fe8B20	86	0.07

Note: The numerical values represent percentages (i.e., Fe19Ni81 means the material contains 19% iron and 81% nickel).

Magnetoresistive sensors usually come in a bridge configuration with four elements in the bridge. This allows adjustment of drift and increases the output from the sensor. As a whole, AMR sensors are very sensitive and can operate at low magnetic fields in a variety of applications, including electronic compasses and magnetic reading heads. Some materials exhibit enhanced magnetoresistive properties. These are often termed giant magnetoresistive (GMR) materials and their use results in GMR sensors with enhanced sensitivities and applications to very low field sensing.

5.7 | MAGNETOSTRICTIVE SENSORS AND ACTUATORS

The magnetostrictive effect is the contraction or expansion of a material under the influence of the magnetic field and the inverse effects of changes in magnetization due to stress in ferromagnetic materials due to the motion of the magnetic walls within the material. This bidirectional effect between the magnetic and mechanical states of a magnetostrictive material is a transduction capability that is used for both actuation and sensing. The effect is an inherent property of some materials. Most materials do not exhibit the effect, but others are strongly magnetostrictive. The effect was first observed in 1842 by James Prescott Joule. There are two effects and their reciprocals as follows:

Effects

1. The **Joule effect** is the change in length of a magnetostrictive sample due to an applied magnetization. This is the most common of the magnetostrictive effects and is quantified by the **magnetostrictive coefficient**, λ, defined as "the fractional change in length as the magnetization of the material increases from zero to its saturation value." Whereas the definition may be obscure to most people, its effects are common: the high-pitched sound emitted by a conventional TV or the humming of a transformer are due to this effect, caused by the magnetization and demagnetization of transformer cores.
2. The reciprocal effect, the change in the susceptibility (i.e., the permeability of the material changes) of a material when subjected to a mechanical stress, is called the **Villari effect**. The change in permeability can be positive or negative. A positive Villari effect is seen when permeability increases, whereas a negative Villari effect means that the permeability decreases.

Reciprocals

1. Twisting of a magnetostrictive sample. When an axial magnetic field is applied to the sample and a current passes through the magnetostrictive sample itself, the interaction between the two causes the twisting effect. This is known as the **Wiedemann effect**, and together with its inverse is used in torque magnetostrictive sensors.
2. The reciprocal effect, that of creation of an axial magnetic field by a magnetostrictive material when subjected to a torque, is known as the **Matteucci effect**.

The magnetostrictive effect is exhibited by the transitional metals, including iron, cobalt, and nickel, and some of their alloys. The magnetostrictive coefficients of some magnetostrictive materials are shown in **Table 5.5**. There are currently materials that exhibit what is called "giant magnetostriction," in which the magnetostrictive coefficient exceeds 1000 μL/L, such as the various metglas (metallic glass) materials and Terfenol-D. These materials are quickly becoming the materials of choice for magnetostrictive sensors and actuators.

The magnetostrictive coefficient is given for the saturation magnetization for each material and hence represents the largest expansion per unit length (or largest strain). For any other magnetic flux density, the strain is proportional. If the saturation magnetostriction is given as in **Table 5.5** at a saturation magnetic flux density B_m, then, assuming linear behavior, the expansion per unit length (strain) at any value of B is

$$\left(\frac{\Delta L}{L}\right)_B = \left(\frac{\Delta L}{L}\right)_{B_m} \times \left(\frac{B}{B_m}\right). \tag{5.59}$$

For a given length L_0, the value above must be multiplied by L_0 to find the absolute expansion of the device.

Some of the earliest uses of magnetostrictive materials date to the beginning of the twentieth century and include telephone receivers, hydrophones, magnetostrictive oscillators, torque meters, and scanning sonar. An early electrical sensor/actuator,

TABLE 5.5 ■ Saturation magnetostriction or magnetostrictive coefficient for some materials

Material	Saturation magnetostriction (magnetostrictive coefficient) [μm/m]	Saturation magnetic flux density [T]
Nickel	−28	0.5
49Co, 49Fe, 2V	−65	
Iron	+5	1.4−1.6
50Ni, 50Fe	+28	
87Fe, 13Al	+30	
95Ni, 5Fe	−35	
Cobalt	−50	0.6
$CoFe_2O_4$	−250	
Galfenol ($Ga_{0.19}Fe_{0.81}$)	50–320	
Terfenol-D ($Tb_{0.3}Dy_{0.7}Fe_2$)	2000	1.0
Vitreloy106A (metglas)(58.2Zr, 15.6Cu, 12.8Ni, 10.3Al, 2.8Nb)	20	1.5
Metglas-2605SC (81Fe, 3.5Si, 13.5B, 2C)	30	1.6

the first telephone receiver (the Reis telephone), tested by Johann Philipp Reis in 1861, was based on magnetostriction.

Applications for magnetostrictive devices include ultrasonic cleaners, high-force linear motors, positioners for adaptive optics, active vibration or noise control systems, medical and industrial ultrasonics, pumps, and sonar. In addition, magnetostrictive linear motors, reaction mass actuators, and tuned vibration absorbers have been designed. Less obvious applications include high cycle accelerated fatigue test stands, mine detection, hearing aids, razor blade sharpeners, and seismic detectors. Ultrasonic magnetostrictive transducers have been developed for surgical tools, underwater sonar, and chemical and material processing.

In general, the magnetostrictive effect is quite small and requires indirect methods for its measurement. There are devices, however, that use the effect directly. The operation of a basic magnetostrictive device is shown in **Figure 5.51**.

There are a number of methods by which magnetostrictive devices may be made to sense a variety of quantities. One of the simplest and most sensitive is to use magnetostrictive materials as the core of simple transformers. This is discussed below in the section on magnetometers. However, most of the applications of magnetostriction are for actuators, which will be described shortly. Nevertheless, indirect use of the magnetostrictive effect can be made to sense a variety of effects, such as position, stress, strain, and torque.

A noncontact magnetostrictive torque sensor for rotating shafts is shown in **Figure 5.52**. It consists of a sleeve of prestressed maraging steel (steel with 18% nickel, 8% cobalt, 5% molybdenum, and small amounts of titanium, copper, aluminum manganese, and silicon) tightly fitted on the shaft itself and two eddy current sensors arranged so that they are at 90° to each other and at 45° to the axis of the shaft, as shown in **Figures 5.52a** and **5.52b**. The torque sensor is based on two principles. First, the magnetostrictive steel, because it is stressed, when it is compressed its permeability decreases, since now the stress is reduced (negative Villari effect). When the steel is tensioned, its stress increases and its permeability increases (positive Villari effect). These can be seen in **Figure 5.52a**, which shows the main compression and tension lines. These are at 45° to the axis of the shaft, hence the choice for orientation of the eddy current sensors (**Figure 5.52b**). The second principle involved is the generation of eddy currents by an AC-driven coil. The eddy currents are influenced by the permeability of the material through the skin effect. A decrease in permeability will cause eddy currents to penetrate deeper into the steel sleeve, whereas an increase in permeability will cause shallower penetration (lower skin depth; see **Equation (5.35)**). Therefore, all that is now necessary is to relate these changes in eddy current induction to the torque responsible for changes in permeability.

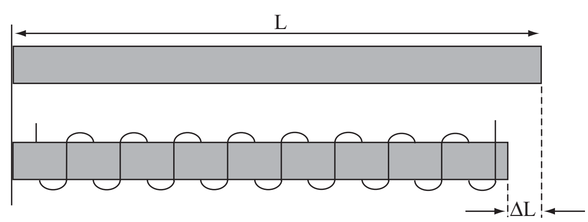


FIGURE 5.51 ■ Operation of a magnetostrictive device.

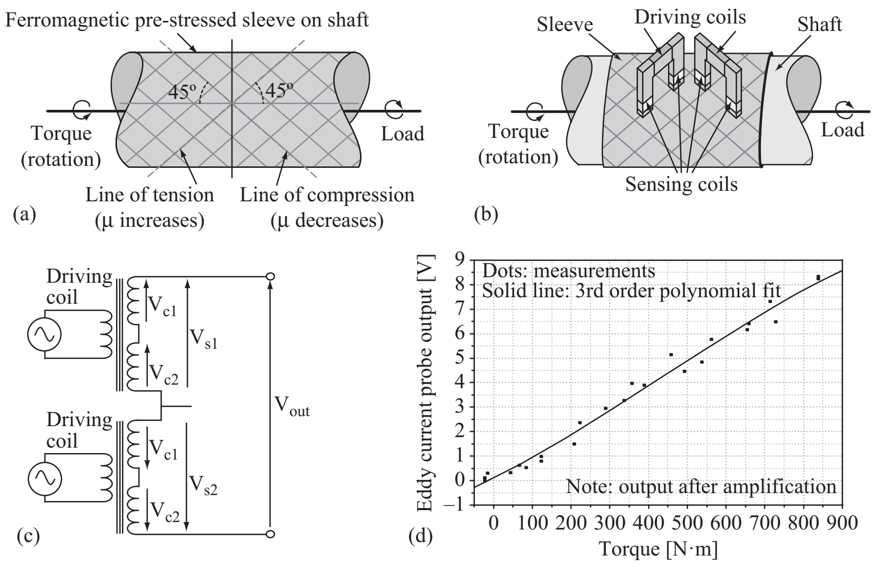


FIGURE 5.52 ■ A magnetostrictive torque sensor. (a) Prestressed ferromagnetic sleeve showing the axes of tension and compression. (b) Two c-core magnetic sensors positioned along the lines of compression and tension form a high-frequency transformer in which the flux closes through the sleeve. (c) The connection of the coils showing the differential output of the device. (d) The transfer function of the torque sensor.

The eddy current sensors consist of a driving coil at the center of the U-shaped core and two pickup coils at its tips. These act as a (high-frequency) transformer, and the voltage induced in the pickup coils depends on the stress condition in the maraging steel ring through which the flux closes (the open ends of the eddy current sensor cores are very close to the rotating sleeve to maximize output, but do not touch). The two driving coils on each sensor are connected in series so that their voltages add up. The two sensing coils are connected in a differential mode, as shown in **Figure 5.52c**.

To understand the operation, suppose first that the torque is zero. Because the two sensors are connected in differential mode, the net output is zero. As the torque increases, one sensor will experience an increase in output while the second will experience a decrease in output. Now the difference between them is the sum of the changes in voltages in the two sensors. Monitoring this voltage (which, of course, can be amplified as needed to produce a convenient span) gives a reading of torque. The experimental output of a sensor of this type is shown in **Figure 5.52d** for a shaft rotating at 200 rpm. As expected, the measurement introduces some errors, and the response of the sensor, although clearly not perfectly linear, is close to a linear response. The solid curve is shown as a polynomial fit of the experimental that can serve as the calibration curve for the sensor.

5.7.1 Magnetostrictive Actuators

Magnetostrictive actuators are quite unique. There are two distinct effects that can be used. One is the constriction (or elongation) or the torque effect produced by the Joule and Wiedemann effects discussed above. The other is due to the stress or shockwave that can be generated when a pulsed magnetic field is applied to a magnetostrictive material.

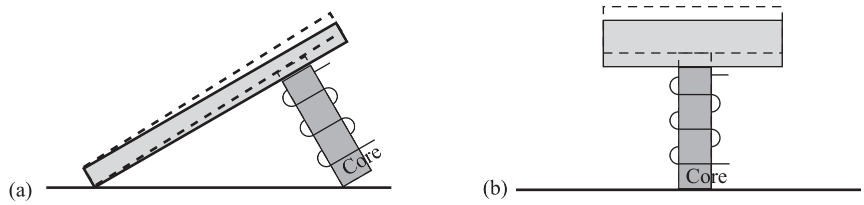
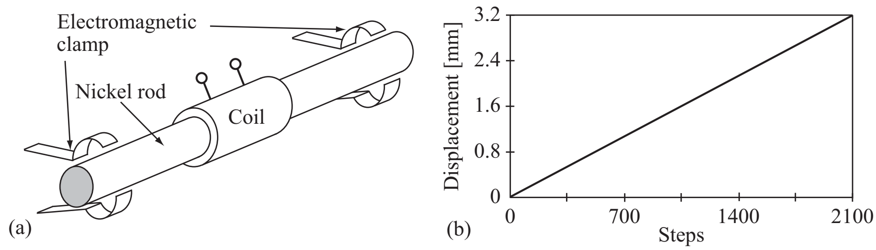


FIGURE 5.53 ■ Micropositioning using a magnetostrictive actuator. (a) Tilting a mirror for optical measurements. (b) Moving or positioning a block. Driving the coil constricts (or expands) the core. Removing the current returns it to its original length.

The first of these is very small (see **Table 5.5**), but it can produce very large forces. It can be used directly in micropositioning (**Figure 5.53**), where very small, accurate, and reversible positioning is possible. This has been used to move micromirrors to deflect light (and in other small structures), but we shall illustrate the idea with the "inchworm" motor shown in **Figure 5.54**. In this device, a bar of nickel is placed between two magnetically actuated clamps. A coil on the bar generates the requisite magnetostriction. By first clamping A, then turning on the current to the coil, end B contracts to the left. Now, clamp B is closed, clamp A is opened, and the current in the coil is turned off. End A elongates back to the original length of the bar and, in effect, the bar has now traveled to the left a distance ΔL, which depends on the magnetostrictive coefficient and the magnetic field in the bar. Whereas the motion in each step is only a few micrometers and motion is necessarily slow, this is a linear motion device that can exert relatively large forces and can be used for accurate positioning. Motion to the right is obtained by reversing the sequence of clamping and current pulses.

FIGURE 5.54 ■ A microstep inchworm motor. (a) Structure. (b) Response.



EXAMPLE 5.10 Linear magnetostrictive actuator

Because magnetostriction is a relatively small effect, many magnetostrictive actuators employ some means of mechanical amplification of the mechanical elongation or constriction of the magnetostrictive element. Consider the actuator in **Figure 5.55a**. The magnetostrictive bar is 30 mm long and is equipped with an elliptical ring or shell whose purpose is to amplify the horizontal magnetostrictive motion to a larger, vertical motion of the shell. The shell also closes much of the magnetic field, requiring a lower current for any given extension of the magnetostrictive bar. The magnetostrictive bar is made of Terfenol-D and the coil can generate magnetic fields varying from 0 to 0.4 T in the bar. Calculate the range of the vertical constriction of the shell.

Solution: Although the horizontal motion of the bar is immediately calculable from **Table 5.5**, we need to do some trigonometric calculations to translate this motion into the vertical motion of the shell. To do so we use **Figure 5.55b**, where we connected the vertical axis point with the horizontal axis point with a line of length l that makes an angle α with the horizontal axis. When the field is applied, the bar elongates (Terfenol-D has a positive magnetostriction coefficient) moving the horizontal axis point to the left and the vertical axis point down. The line now makes an angle β with the horizontal axis, but its length remains the same. From these two angles and the horizontal and vertical displacements we write

$$l\cos\beta = l\cos\alpha + \Delta x, \quad l\sin\beta = l\sin\alpha - \Delta y.$$

From the second relation,

$$\sin\beta = \frac{l\sin\alpha - \Delta y}{l} \rightarrow \cos\beta = \sqrt{1 - \sin^2\beta} = \frac{1}{l}\sqrt{l^2 - (l\sin\alpha - \Delta y)^2},$$

and substituting this into the first relation,

$$\sqrt{l^2 - (l\sin\alpha - \Delta y)^2} = l\cos\alpha + \Delta x.$$

Squaring both sides and arranging terms gives:

$$l^2 - (l\sin\alpha - \Delta y)^2 = (l\cos\alpha + \Delta x)^2 \rightarrow \Delta y^2 - 2l\sin\alpha\Delta y + \Delta x^2 - 2l\sin\alpha\Delta x$$
$$= l^2 - l^2\sin^2\alpha - l^2\cos^2\alpha.$$

The right-hand side of the relation is zero since $\sin^2\alpha + \cos^2\alpha = 1$ and we have a second-order equation with the unknown variable Δy in terms of known variables Δx, l, and α. Solving and taking only the positive root, we get

$$\Delta y = l\sin\alpha - \sqrt{l^2\sin^2\alpha - \Delta x(\Delta x + 2l\cos\alpha)} \quad [\text{m}].$$

In the case in **Figure 5.55a**, the angle α is

$$\alpha = \tan^{-1}\left(\frac{7.5}{15}\right) = 26.565°$$

and the length l is

$$\frac{15}{l} = \cos 26.565° \rightarrow l = \frac{15}{\cos 26.565°} = 16.77 \text{ mm}.$$

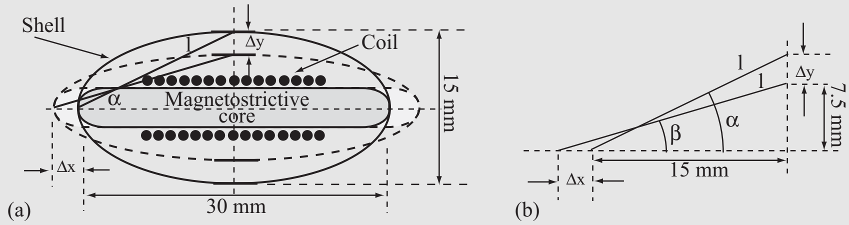


FIGURE 5.55 ■ A practical magnetostrictive actuator with mechanical amplification. (a) Structure and operation. (b) Calculation of displacement.

Now we can calculate Δx from **Table 5.5**. Since $\Delta l/l = 2000$ at a saturation flux density of 1 T, for a bar of length 30 mm and a flux density of 0.4 T we have a total expansion of

$$\Delta l = \left(\frac{\Delta l}{l}\right)_{B_m} \times \left(\frac{B}{B_m}\right) l = 2000 \times 10^{-6} \times \left(\frac{0.4}{1.0}\right) \times 30 = 2.4 \text{ mm}.$$

Now, since $\Delta l = 2\Delta x$ in **Figure 5.55b**, we can calculate Δy from the above:

$$\Delta y = l \sin\alpha - \sqrt{l^2 \sin^2\alpha - \Delta x(\Delta x + 2l\cos\alpha)} = 16.77 \sin 26.565^\circ$$
$$- \sqrt{(16.77)^2 \times \sin^2 26.565^\circ - 1.2 \times (1.2 + 2 \times 16.77 \times \cos 26.565^\circ)} = 3.163 \text{ mm}.$$

The shell moves twice as far since each side moves the same distance. The total motion is therefore 6.326 mm. In effect, the structure has generated an amplification of 6.326/2.4 = 2.64.

5.8 | MAGNETOMETERS

In general, magnetometers are devices that measure magnetic fields and, as such, the name can be assigned to almost any system that can measure the magnetic field. However, properly used, it refers to either very accurate sensors or low field sensing on the one hand or, on the other hand, complete systems for measuring the magnetic field, which include one or more sensors. We shall use the term as a sensor for low field sensing since it is in this capacity that magnetometers become unique devices. There are a number of methods, sometimes unrelated to each other.

5.8.1 Coil Magnetometer

To understand the fundamental methods of sensing we will start with the simplest idea of sensing a field—the small coil shown in **Figure 5.56**. In this coil the emf (voltage) measured across the coil is

$$emf = -N\frac{d\Phi}{dt} \quad [\text{V}], \text{ where } \Phi = \int_S BS \sin\theta_{BS} \quad [\text{Wb}], \tag{5.60}$$

where Φ is the flux through the coil, N is the number of turns in the coil, and θ_{BS} is the angle between the direction of the field and the normal to the area of the coil. This

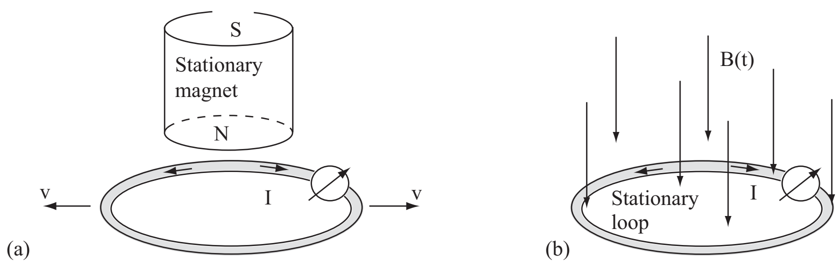


FIGURE 5.56 ■ Operation of a small search coil. (a) The coil moves in a constant magnet field. Induced emf (and hence current) depends on the motion of the coil in the DC magnetic field. (b) Induced current in a stationary coil by a time-dependent magnetic field.

relation is called Faraday's law of induction. The relations show that the output is integrating (dependent on the coil's area). This basic device indicates that to measure local fields, the area of the coil must be small, that sensitivity depends on the size and number of turns, the frequency, and, from **Figure 5.56**, that only variations in the field (due to motion in **Figure 5.56a** or due to the AC nature of the field in **Figure 5.56b**) can be detected. If the field is time dependent, it can be detected with stationary coils as well.

There are many variations on this basic device. First, differential coils may be used to detect spatial variations in the field. In other magnetometers, the coil's emf is not measured. Rather, the coil is part of an *LC* oscillator and the frequency is then inductance dependent. Any conducting and/or ferromagnetic material will alter the inductance and hence the frequency. This creates a relatively sensitive magnetometer often used in areas such as mine detection or buried object detectors (pipe detection, "treasure" hunting, etc.). While the simple coil, in all its configurations, is not normally considered a particularly sensitive magnetometer, it is often used because of its simplicity and, if properly designed and used, can be reasonably sensitive (e.g., magnetometers based on two coils have been used for airborne magnetic surveillance for mineral exploration and for detection of submarines).

EXAMPLE 5.11 The coil magnetometer

A magnetometer of surprising sensitivity can be made with a simple coil. Consider a device intended to detect and trace current-carrying wires buried in a wall or to map the magnetic field in a house or in the vicinity of power lines. The sensor itself is a simple coil, as shown in **Figure 5.57**. The emf produced by the coil based on **Equation (5.60)** is amplified and displayed or an alarm is sounded. The coil given here has 1000 turns and an average diameter of 4 cm and we wish to calculate its output due to AC magnetic fields. Assuming that a minimum output of 20 mV is needed to overcome background noise, what is the lowest magnetic flux density produced by a power line at 60 Hz that can be measured?

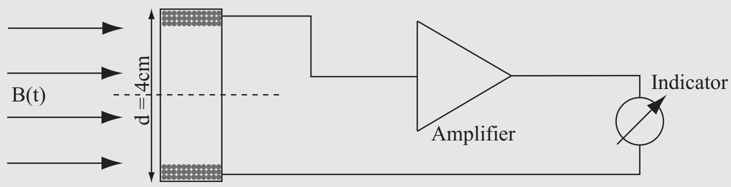


FIGURE 5.57 ■ The coil magnetometer. The amplifier is often needed unless the field is high.

Solution: We will assume that the magnetic flux density is uniform across the plane of the coil, and since the flux density is sinusoidal we write

$$emf = -N\frac{d\Phi(t)}{dt} = -NS\frac{dB(t)}{dt} = -NS\frac{d}{dt}B\sin(2\pi ft) = -2\pi fNSB\cos(2\pi ft) \quad [\mathrm{V}].$$

For a 20 mV emf we have (neglecting the negative sign, which only indicates the phase of the emf relative to the magnetic flux density)

$$B = \frac{emf}{2\pi fNS} = \frac{20\times10^{-3}}{2\pi\times60\times1000\times\pi(2\times10^{-2})^2} = 4.22\times10^{-5}\ \mathrm{T}.$$

This is on the same order of magnitude as the terrestrial magnetic flux density (approximately 60 μT).

Notes:

1. The device can only detect time dependent fields.
2. The sensitivity can be increased by increasing the number of turns, the dimensions of the coil, or if the frequency is higher. A ferromagnetic core in the coil can also help by concentrating the flux density in the coil.
3. The principle outlined here is the basis of many simple magnetometers or “field testers” ranging from “live wire detectors” in walls to “gauss meters” for general field measurement use, including electromagnetic compatibility tests and the detection of minerals ‘by detecting variations in the terrestrial magnetic field using large differential magnetometers.

5.8.2 The Fluxgate Magnetometer

Much more sensitive magnetometers can be built on the basis of fluxgate sensors. The fluxgate sensor can also be used as a general purpose magnetic sensor, but it is more complex than the simple sensors described above such as the magnetoresistive sensor. It is therefore most often used where other magnetic sensors are not sensitive enough. Examples are electronic compasses, detection of fields produced by the human heart, or fields in space. These sensors have existed for many decades, but are rather large, bulky, and complex instruments specifically built for applications in scientific research. Recently they have become available as off-the-shelf sensors because of developments in new magnetostrictive materials that has allowed their miniaturization and even integration in hybrid semiconductor circuits. New fabrication techniques promise to improve these in the future, and as their size decreases, their uses will expand.

The idea of a fluxgate sensor is shown in **Figure 5.58a**. The basic principle is to compare the drive-coil current needed to saturate the core in one direction to that needed to saturate it in the opposite direction. The difference is due to the external field. In practice, it is not necessary to saturate the core, but rather to bring the core into its nonlinear range (see **Figure 5.19**). The magnetization curve for most ferromagnetic materials is highly nonlinear, so almost any ferromagnetic material is suitable. In practice, the coil is driven with an AC source (sinusoidal or square, but more often a triangular source), and under no external field, the magnetization is uniform throughout the coil and hence the sense coil will produce zero output. An external magnetic field

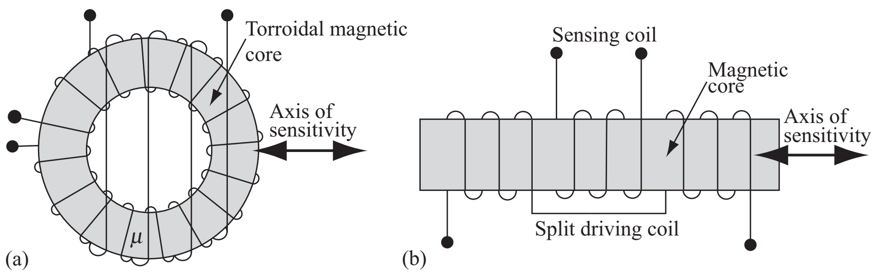


FIGURE 5.58 ■ Principle of the fluxgate magnetometer. (a) Toroidal implementation; the axis of sensitivity is shown. (b) Bar or film implementation. Sensitivity is along the long axis of the bar.

perpendicular to the sense coil changes the magnetization and, in effect, the core has now become nonuniformly magnetized, producing an emf in the sensing coil on the order of a few millivolts per microteslas [mV/μT]. The reason for the name fluxgate is this switching of the flux in the core in opposite directions. The same can be achieved by using a simple rod, as in **Figure 5.58b**. Now the two coils are wound one on top of the other and the device is sensitive to fields in the direction of the rod, but the principle is the same, the output relies on variations in permeability (nonlinearity) along the bar. A particularly useful configuration is the use of a magnetostrictive film (metglasses are a common choice), as shown in **Figure 5.58b** or a similar configuration. Because magnetostrictive materials are highly nonlinear, the sensors so produced are extremely sensitive—with sensitivities of 10^{-6} to 10^{-9} T. The sensors can be designed with two or three axes. For example, in **Figure 5.58a**, a second sensing coil can be wound perpendicular to the first. This coil will be sensitive to fields perpendicular to its area and the whole sensor now becomes a two-axis sensor. Fluxgate sensors are available in integrated circuits where permalloy is the choice material, since it can be deposited in thin films and its saturation field is low. Nevertheless, current integrated fluxgate sensors have lower sensitivities than the classical fluxgate sensors—on the order of 100 μT—but still higher than many other magnetic field sensors.

To better understand the operation of the fluxgate sensor, consider one particular and highly useful form, the so-called pulse-position fluxgate sensor. In this type of sensor, the current in the driving coil is triangular and produces a likewise triangular-shaped flux in the core. This flux may be considered a reference flux. An additional flux is produced by the external magnetic field—the field being measured. This flux either adds or subtracts from the internal, reference flux, but in any case, it produces an emf in the sensing coil. This is the measured flux. The two emfs—one from the reference flux and the other from the measured flux—are then compared. Whenever the reference field is higher than the measured field, the output is high, and whenever it is lower, the output is zero (we will see in **Chapter 11** that this can be accomplished by a simple circuit called a comparator). **Example 5.12** discusses this sensor and simulates an implementation to predict its sensitivity prior to construction.

EXAMPLE 5.12 The pulse-position fluxgate sensor: simulation results

To appreciate the value of simulation in design and to understand the operation of the fluxgate sensor, consider a fluxgate sensor as in **Figure 5.58a**, based on the pulse-position principle. The sensor and the circuits were simulated first and the results shown here are those of the simulation. This allows one to obtain results free of noise and only then build the actual circuit (shown in **Figure 5.60**). To do so, the properties of the core, the coils, and the electronic circuits are inserted into the simulation and appropriate signals are generated. In this case, a toroidal core with inner radius of 19.55 mm, outer radius of 39.25 mm, and cross-sectional area of 231 mm^2 with a saturation flux density of 0.35 T and relative permeability of 5000 is simulated. The core is wound with 12 turns for the drive coil and 50 turns for the sensing coil. The triangular reference signal oscillates at 2.5 kHz between −4.5 V and +4.5 V. The simulation produces any required number of results. We show three. **Figure 5.59a** shows the flux density in the core produced by the internal triangular generator. It shows a symmetric flux density ranging from +1 mT to −1 mT. **Figure 5.59b** shows the flux density in the core when an external flux density B_e of 500 μT

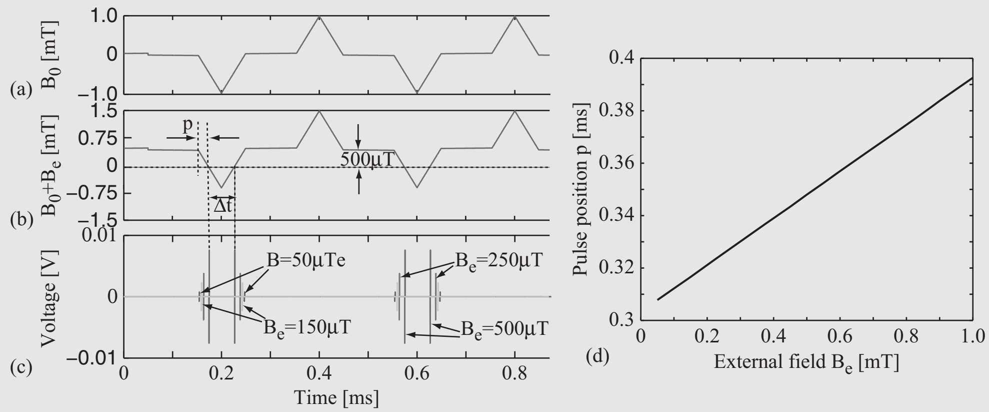


FIGURE 5.59 ■ (a) Magnetic flux density generated in the core using a triangular pulse generator. (b) The magnetic flux density in the core when an external flux density of 500 μT is measured. (c) The pulse positions obtained by detecting the zero-crossings. (d) Transfer function of the sensor given as the pulse position p versus the external magnetic flux density B_e.

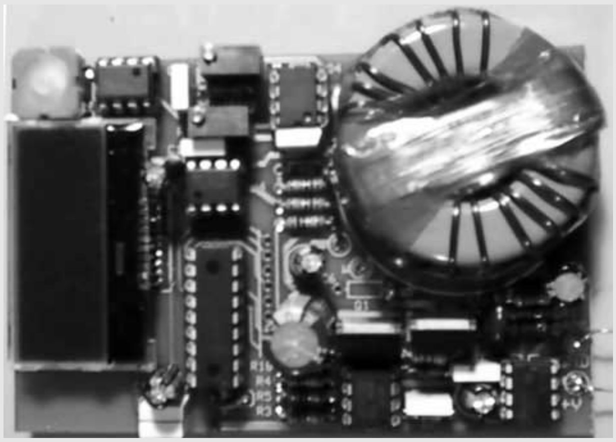


FIGURE 5.60 ■ Implementation of the fluxgate sensor. Note the toroidal coil at the upper right corner. In this implementation, based on a microprocessor, the flux density is displayed numerically on the display on the left.

is sensed. The flux density in **Figure 5.59b** is the sum of the reference (drive) flux density and the measured flux density, as can be seen by the shift upward of the zero line. The device detects the zero crossing of the flux density. The result is the output pulse in **Figure 5.59c**, showing the time width between the two comparison points. The pulse position can be taken as the distance between the pulses shown in **Figure 5.59c** and the position of the pulse (zero-crossing) when no field is applied, indicated in **Figure 5.59b** as p. Alternatively one can measure the time Δt shown in **Figure 5.59b**. **Figure 5.59d** shows the simulated transfer function indicating a linear relation and a sensitivity of 88 ms/T. This may not seem very large, but considering the fact that a time of 1 μs can be easily and accurately measured, the device can easily measure about 10 nT. The device implemented and tested based on the simulations described in this example is shown in **Figure 5.60**.

5.8.3 The SQUID

SQUID stands for superconducting quantum interference device. By far the most sensitive of all magnetometers, they can sense down to 10^{-15} T, but this kind of performance comes at a price—they operate at very low temperatures, usually at 4.2°K (liquid helium). As such, they do not seem to be the type of sensor one can simply take off the shelf and use. Surprisingly, however, higher temperature SQUIDs and integrated SQUIDs do exist. Even so, they are not as common as other types of sensors. The reason for including them here is that they represent the limits of sensing and have specific applications in the sensing of biomagnetic fields and in testing materials integrity.

The SQUID is based on the Josephson junction, formed by two superconductors separated by a small insulating gap (discovered in 1962 by Brian David Josephson). If the insulator between two superconductors is thin enough, the superconducting electrons can tunnel through it. For this purpose the most common junction is the oxide junction in a semiconductor, but there are other types. The base material is usually niobium or a lead (90%)–gold (10%) alloy with the oxide layer formed on small electrodes made of the base material that are then sandwiched to form the junction.

There are two basic types of SQUIDs. Radio frequency (RF) SQUIDs, which have only one Josephson junction and DC SQUIDs, which usually have two junctions. DC SQUIDs are more expensive to produce, but are much more sensitive.

If two Josephson junctions are connected in parallel (in a loop), electrons, which tunnel through the junctions, interfere with one another. This is caused by a phase difference between the quantum mechanical wave functions of the electrons, which is dependent upon the strength of the magnetic field through the loop. The resultant superconducting current varies with any externally applied magnetic field. The external magnetic field causes a modulation of the superconducting current through the loop, which can be measured (**Figure 5.61**). The superconducting current is set up externally by the sense loop (a single loop as in **Figure 5.61a** is used to measure fields, while two loops as in **Figure 5.61b** are used to measure the gradient in the field). The current may also be set up directly by the superconducting loop. The output is the change in voltage across the junction due to changes in the current, and since the junction is resistive, this change is measurable following amplification.

An RF SQUID operates in the same fashion except that there is only one junction and the loop is driven by an external resonant circuit that oscillates at high frequency (20–30 MHz). Any change in the internal state of the flux in the measurement loop (due to external, sensed fields) changes the resonant frequency, which is then detected and is a measure of the field.

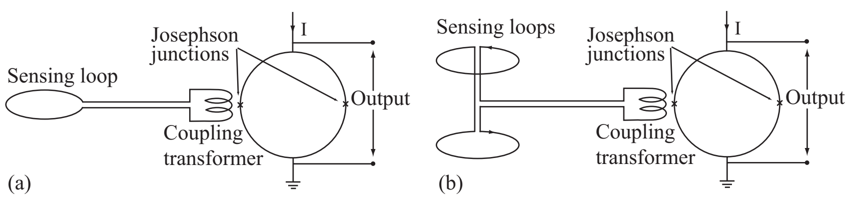


FIGURE 5.61 ■ SQUID structure and operation. (a) Measurement of magnetic fields. (b) Measurement of gradients in the magnetic field.

The main difficulty with SQUIDs is the cooling needed and the necessary bulk. Nevertheless, it is an exceedingly useful sensor where and when the cost can be justified. It is exclusively used in applications such as magnetoencephalography (measurement of magnetic fields of the brain) and in research. Many applications require that measurements of very low magnetic fields be done in shielded rooms where the terrestrial magnetic field as well as other sources, such as those from power lines, can be eliminated.

5.9 | MAGNETIC ACTUATORS

We have already discussed magnetostrictive actuators, but there are many other types of magnetic actuators and many of them are more conventional. In particular, the whole group of electric motors forms the bulk of conventional (and some less than conventional) actuation applications. The main reason that magnetic actuation is usually the best choice has to do with energy density, just as the reason for the use of the internal combustion engine has to do with energy density. An example is instructive. If we restrict ourselves to electrical actuation, there are two basic types of forces that exist. One is the Coulomb force, which acts on a charge in the electric field (see **Section 5.3.3**). The second is the magnetic force, which acts on a current in a magnetic field as defined by the Lorentz force in **Section 5.4**. These lead to energy densities that are given as follows:

Electric energy density:

$$w_e = \frac{\varepsilon E^2}{2} \quad [\mathrm{J/m^3}] \tag{5.61}$$

Magnetic energy density:

$$w_m = \frac{B^2}{2\mu} \quad [\mathrm{J/m^3}], \tag{5.62}$$

where E is the electric field intensity and B is the magnetic flux density. Now, by way of comparison, suppose we use a common value for the magnetic flux density B in an electric motor of 1 T. The permeability of iron in the motor is about $1000\mu_0 = 1000 \times 4\pi \times 10^{-7}$ H/m. This gives an energy density in the motor of

$$w_m = \frac{1^2}{2 \times 1000 \times 4\pi \times 10^{-7}} \approx 400\ \mathrm{J/m^3}. \tag{5.63}$$

On the other hand, the relative permittivity of most common dielectrics is less than 10. Taking an electric field intensity of 10^5 V/m (this is a very large electric field intensity) and a permittivity of $10\varepsilon_0 = 8.845 \times 10^{-11}$ F/m, the energy density in a purely electric actuator becomes

$$w_e = 8.854 \times 10^{-11}\,\frac{10^{10}}{2} \approx 0.45\ \mathrm{J/m^3}. \tag{5.64}$$

Note: In most cases the absolute maximum electric field intensity cannot exceed a few million volts per meter without breakdown, so increasing the electric field to its absolute maximum will only increase the energy density by about one order of magnitude.

Clearly then, a magnetic actuator will be capable of exerting larger forces in a smaller package than an electric actuator. In addition, it is usually easier and much safer to generate fairly large magnetic fields than to generate large electric fields. Nevertheless, it should be mentioned that electric forces and electric actuators have their own niche application in MEMS, where the required forces are small and small energy densities are acceptable (see **Chapter 10**). They are also useful in electrostatic filters and dust collectors, where the small forces are sufficient to act on charged dust particles and collect them on electrodes. Electrostatic actuation is also critical to the operation of copiers and printers, where toner particles are distributed on a printing medium using electrostatic forces.

But the inevitable consequence is that most electric actuators are based on magnetic forces, and among these, motors feature prominently. However, there are many types of motors and related devices, such as voice coil actuators and solenoids. Motors and solenoids will be discussed at length, but we start with a particular type of magnetic actuator called the voice coil actuator, because the discussion also introduces the principle upon which motors are based.

5.9.1 Voice Coil Actuators

Voice coil actuators got their name from their first and perhaps still their most widely used implementation—that of magnetically driven loudspeakers. In most applications of voice coil actuators, there is no use of voice—only the similarity in operation. The actuators are based on the interaction between the current in a coil and the magnetic field of a permanent magnet or another coil. To understand this, consider the basic structure of a loudspeaker mechanism shown in **Figure 5.62** (loudspeakers will be discussed in their own right in **Chapter 7**). The magnetic field in the gap is radial. For a current-carrying loop, the force is given by **Equation (5.26)** (Lorentz force), where L is the circumference of the loop and we assume a uniform magnetic field. With N turns, the force is $NBIL$. Of course, the field does not have to be uniform or the coil circular, but this is a simple configuration and is the one used in most speakers. The larger the current, the larger the force, and thus the larger the displacement of the speaker's cone. By reversing the current, the coil moves in the opposite direction. Before we proceed we should note the following:

1. The force is directly proportional to current for a given magnetic field. In this case (and in many voice coil actuators) it is linear with current, an important property of the device.

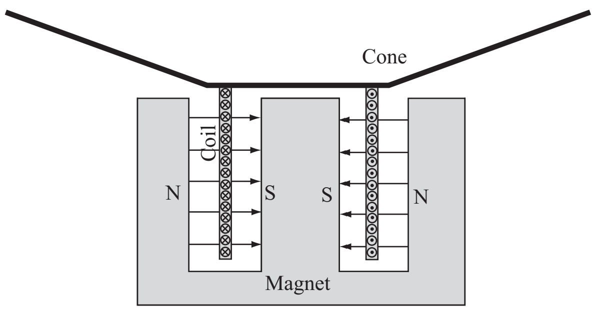
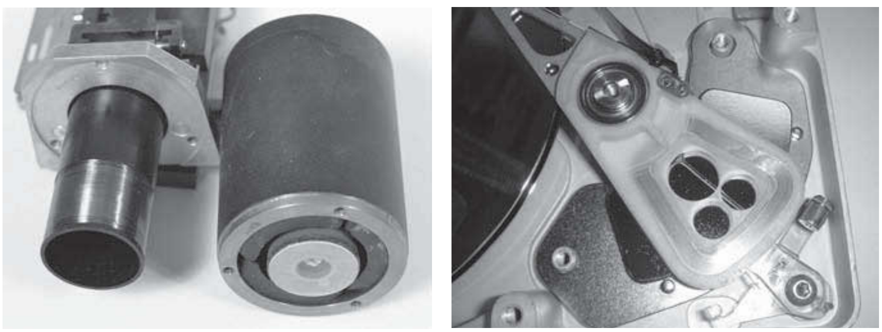


FIGURE 5.62 ■ The structure of a loudspeaker showing the interaction between the magnetic field and current in the coil.

2. The larger the coil or the magnetic field, the larger the force.
3. By allowing the coil to move, the displaced mass is small (compared with other actuators) and hence the mechanical response is rather good. For this reason, a speaker can operate at, say, 15 kHz, whereas a motor-driven actuator may take seconds to reverse.
4. It is also possible to fix the coil and allow the magnet to move.
5. The field in the actuator can be generated by an electromagnet if necessary.
6. We note here that the voice coil actuator can be turned into a sensor by simply reversing the action. If we move the coil in the magnetic field, the voltage induced in the coil will be given by Faraday's law of induction through **Equation (5.60)**. The speaker becomes a microphone and the more general voice coil actuator becomes a sensor.
7. In the absence of current, the actuator is entirely disengaged—there is no intrinsic retaining or cogging force and no friction. However, in some cases use may be made of a restoring spring to return the actuator to its rest position, as is done in speakers.
8. The motion is limited and often quite short.
9. Rotational motion can also be achieved by selecting particular coil and magnet configurations (see **Figure 5.63b**).
10. The actuator is a direct drive device.

Among these properties, the main quality that has made voice coil actuators so appealing beyond their use in speakers is that their small mass allows very high accelerations (upwards of 50 g and for very short strokes up to 300 g) at high frequencies, making them ideal candidates for fast positioning systems (e.g., in positioning read/write heads in disk drives). The forces achievable are modest in comparison to other motors, but are certainly not negligible (up to 5000 N) and the power they can handle is also significant.

Voice coil actuators are often used where very accurate positioning at high speeds is needed. Since they have no hysteresis and minimal friction, they are extremely accurate both as linear and as angular positioners. No other actuator matches their response and



(a) (b)

FIGURE 5.63 ■ (a) A cylindrical linear voice coil actuator shown disassembled. The magnets and the gap in which the coil assembly fits are visible on the front of the assembly. (b) Rotary voice coil actuator used as a head positioner in a disk drive. Two permanent magnets can be seen under the trapezoidal coil. A steel cover closes the magnetic circuit above the coil, but it has been removed to show the coil.

acceleration. They can also be used in less critical applications, mostly in positioning and control, and also in valve actuation, pumps, and the like. Interfacing with microprocessors is usually simpler than other types of motors, and control and feedback are easily incorporated.

A large variety of voice coil actuators are available, but the cylindrical actuator in **Figure 5.63a** and the rotary actuator in **Figure 5.63b** are typical. In the cylindrical linear actuator, the magnetic field is radial, as in the loudspeaker. The coil, attached to the moving, actuating shaft, moves in and out from a center position with a maximum stroke defined by the length of the coil and the length of the cylindrical magnet. For motion to be linearly proportional to current, the coil must be within the uniform magnetic field during the entire range of motion. Ratings of these actuators are in terms of stroke, force (in newtons), acceleration, and power.

EXAMPLE 5.13 Force and acceleration a in voice coil actuator

Consider the voice coil actuator shown in **Figure 5.64a**. The coil is wound on a plastic form that rides on the inner core. Assume the coil has 400 turns, the current needed for operation is 200 mA, and that the coil never leaves the area of uniform field. The field itself is produced by a permanent magnet and equals 0.6 T anywhere within the space occupied by the moving part of the actuator. The given dimensions are $a = 2$ mm, $b = 40$ mm, and $c = 20$ mm. Calculate the force and acceleration of the actuator if the moving part weighs 80 g.

Solution: We will use **Equation (5.26)** since it gives the force on a length of wire, but we will modify it slightly to fit this configuration. Since the length of the loops vary with position in the coil (i.e., the length is $2\pi r$, where r is the radius of the loop), we first calculate a current density in the cross section of the coil by multiplying the current by the number of turns and then dividing by the cross-sectional area of the coil:

$$J = \frac{NI}{ab} = \frac{400 \times 0.2}{0.002 \times 0.04} = 1 \times 10^6 \ \mathrm{A/m^2}.$$

Now we define a ring of current of thickness dr and radius r as shown in **Figure 5.64b**. The total current in this ring is

$$dI = Jds = Jbdr \quad [\mathrm{A}].$$

The length of the current ring is taken as its circumference since it is on this length of the current that the magnetic flux density acts. The force on this ring of current is

$$dF = BLdI = B(2\pi r)Jbdr \quad [\mathrm{N}].$$

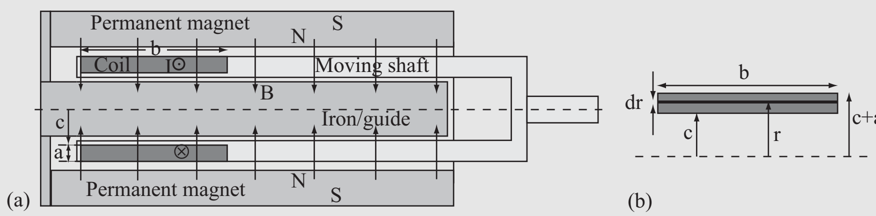


FIGURE 5.64 ■ (a) A cylindrical voice coil actuator. (b) Calculation of force on the coil.

Thus the total force on the coil is

$$F = 2\pi BJb \int_{r=c}^{c+a} r dr = 2\pi BJb \left[\frac{r^2}{2}\right]_{r=c}^{r=c+a} = \pi BJb\left[(c+a)^2 - a^2\right] \quad [\mathrm{N}].$$

Numerically this is

$$F = \pi BJb\left[(c+a)^2 - a^2\right] = \pi \times 0.6 \times 1 \times 10^6 \times 0.04\left[(0.04 + 0.002)^2 - 0.002^2\right] = 120.64\ \mathrm{N}.$$

The acceleration is calculated from Newton's law:

$$F = ma \rightarrow a = \frac{F}{m} = \frac{120.64}{0.080} = 1507.96\ \mathrm{m/s^2}.$$

With a force of 120.64 N and an acceleration of just over 150 g, this is a respectable actuator ($g = 9.81\ \mathrm{m/s^2}$).

Note: The force can be increased by increasing the radius of the moving coil, by increasing the number of turns, and/or by increasing the current in the coil. On the other hand, increasing the number of turns and physical dimensions increases the mass of the coil and hence tends to reduce acceleration, and necessarily reduces the response time of the actuator.

5.9.2 Motors as Actuators

The most common of all actuators are electrical motors in their many types and variations. It would be totally presumptuous to even try to discuss all motors, their principles, and their applications here—many volumes have been used to do so. It is important, however, to discuss some of the more salient issues associated with their use as actuators. Also, at the outset, it should be understood that motors can be used, and often are, as sensors. In fact, many motors can be used as generators, in which capacity they can sense motion, rotation, linear and angular position, and any other quantities that affect these, such as wind speed, flow velocity and rate, and much more. Some of these sensor applications will be discussed throughout this text (e.g., wind speed is commonly sensed and measured by a rotating cup arrangement that turns a small motor and the output is then proportional to wind speed), but for now we shall concentrate on their use as actuators.

Also to be recognized at this stage is that most motors are magnetic devices—they operate by attraction or repulsion between current-carrying conductors or between current-carrying conductors and permanent magnets in a manner similar to that of voice coil actuators. Unlike voice coil actuators, motors include magnetic materials (mostly iron), in addition to permanent magnets or electromagnets, to increase and concentrate the magnetic flux density and in so doing increasing power, efficiency, and available torque in the smallest possible volume. The variation in size, and power, they can deliver is staggering. Some motors are truly tiny. For example, the motors used as vibrators in cell phones are about 6–8 mm in diameter and no more than 20 mm long—some are flat, the size of a small button. On the other end, motors delivering hundreds of megawatts of power are used in the steel and mining industries. Perhaps the largest are generators in power plants—these can be 1000 MW or more. Yet there is no fundamental difference in operation between these devices.

Motors, as their name implies, deliver motion in one form or another. As such, many devices can be called motors. For example, the windup spring mechanism in a clock is a true motor.

As a general classification for actuation purposes, there are three types of motors: continuous rotational motors, stepper motors, and linear motors. Of these, the best known to the casual observer is the continuous rotational motor. However, stepper motors are much more common than one realizes, and linear motors, while not as common, are increasingly finding application in specialized systems. In a continuous motor, the shaft rotates in one direction as long as power is supplied to the motor. Some motors can be reversed (such as DC motors), while some cannot. Stepper motors provide discrete motion or steps of predetermined size. A pulse given to the motor will move it a fraction of a rotation (1°–5° is typical). To move it further, an additional pulse is necessary. This is a boon to positioning, where the accurate and repeatable steps of these motors is very useful. Linear motors are somewhere in between. First, their motion is linear rather than rotational. Because of this, they cannot be truly continuous and must be reversible. Often they are stepping motors, but they can be continuous in the range of their linear motion.

When used in actuation, some form of control is often necessary. One may need to control the speed, direction of motion, number of steps, torque applied, etc. These controls are often accomplished with microprocessors and interfacing circuits (**Chapter 12**) and become part of the actuation strategy for the system. An important class of motors is the so-called brushless DC (BLDC) motors, which is somewhere between continuous and stepping motors. Its importance comes from its control and, as a consequence, from its use in a plethora of applications, especially in data storage drives, CD drives, and in some rotary tools and toys, such as toy airplanes.

In the following sections we shall discuss, briefly, some of the important properties of motors, such as torque, power, speed, etc., and common types of motors used for actuation. Although no specific power range is implied, it should be understood that very large motors have special requirements that small motors do not (mechanical structures, power supplies, cooling, etc.). Therefore the discussion here should be viewed as relating to low-power motors.

5.9.2.1 Operation Principles

All motors operate on the principle of repulsion or attraction between magnetic poles. As an initial discussion, consider the device in **Figure 5.65a**. The two magnets are kept

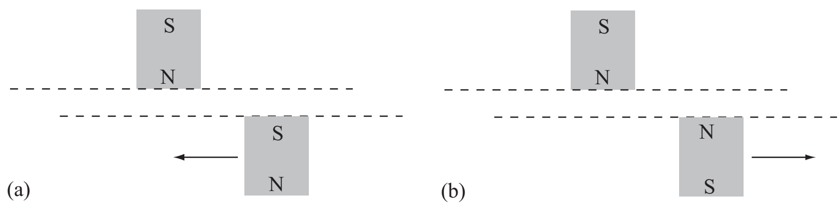


FIGURE 5.65 ■ Forces between two magnetic poles. In this schematic, the upper magnet is fixed (stator) and the lower is free to move (rotor). (a) Attraction. (b) Repulsion.

separated vertically, but the lower magnet is free to move horizontally. The two opposite poles attract and the lower magnet will move to the left until it is aligned with the upper magnet. **Figure 5.65b** is similar, but now the two magnets repel each other and the lower magnet will move to the right. In this very simple example, this is the extent of motion, but this is the first and most fundamental principle of motors. As a matter of nomenclature, the stationary pole is called a stator, whereas the moving pole is the rotor (or in the case of linear motors, a slider).

To make this into a more useful device, consider **Figure 5.66a**. Here the configuration is somewhat different, but the principle is still the same. The magnetic field (which may be produced by a permanent magnet or an electromagnet) is assumed to be constant in time and space for now. If we apply a current to the loop, and assuming the loop is initially at an angle to the field, as shown in **Figure 5.66b**, a force will exist on each of the upper and lower members of the loop equal to $F = BIh$ (Lorentz force in **Equation (5.26)**, where B is the magnetic flux density, I is the current, and h is the length of the loop). This force will rotate the loop to the right one-half turn, until the loop is perpendicular to the magnetic field. Note that the Lorentz force is always perpendicular to both the current and the magnetic field. For a motor to operate continuously, when the loop reaches this position, the current in the loop is reversed (commutated) and, assuming that the loop rotates slightly past the perpendicular position due to inertia, the force now will continue rotating it clockwise an additional half turn, and so on. Note also that the force on the loop is constant (independent of position). Obviously some additional issues have to be resolved, otherwise the loop can get stuck into a vertical position. However, without complicating the issue, it is obvious from this configuration that the motor develops a force and the force, acting on the loop, develops a torque. The latter is

$$T = 2BIhr\sin\alpha \quad [\mathrm{N}\cdot\mathrm{m}], \tag{5.65}$$

where r is the radius of the loop. If multiple loops are used, the force, and hence the torque, is multiplied by the number of loops, N. This particular configuration requires commutation, and this can be done mechanically or electronically. **Figure 5.67a** shows the same configuration with a mechanical commutator and a permanent magnet stator producing the magnetic field. This is a simple DC motor. As the loop rotates, the commutator rotates with it. The contact with the brushes reverses the current at the right moment at the end of each turn. The number of coils can be increased, say, to two as in

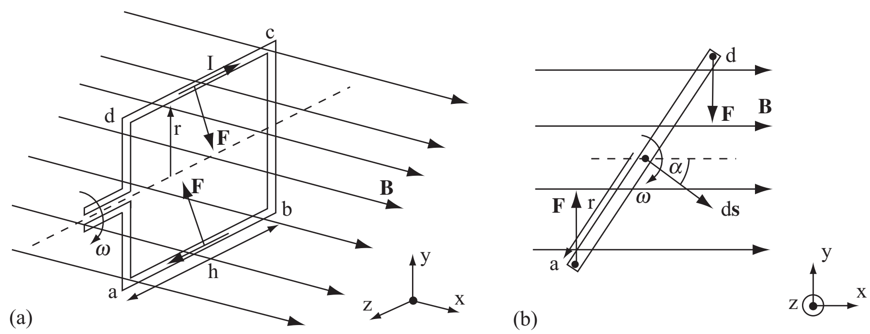


FIGURE 5.66 ■ A loop in a magnetic field showing the force on the current in the loop. (a) The force rotates the loop to the right until the loop's area is perpendicular to the field. (b) Relation between force and position of the loop.

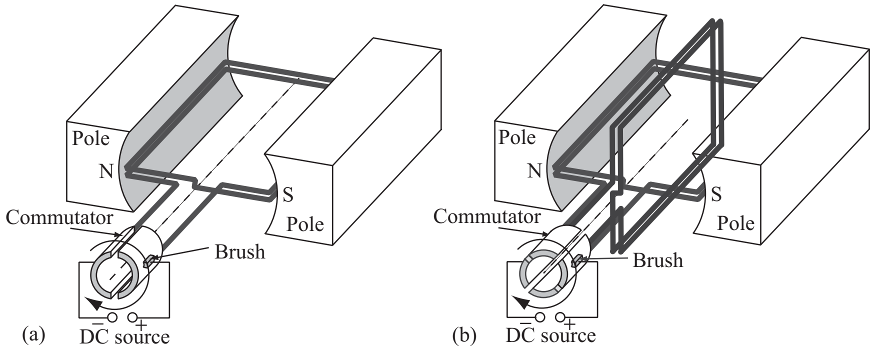


FIGURE 5.67 ■ A commutated DC motor. (a) Single coil with two commutating connections. (b) Two perpendicular coils, each with a pair of commutating connections. The brushes are stationary.

Figure 5.67b. In this case, there are four connections on the commutator so that each coil is powered in the appropriate sequence to ensure continuous rotation. In practical motors of this type, many more coils are used spaced equally around the circumference accompanied by additional commutator connections. This increases torque and makes for smoother operation due to commutation. Most small DC motors are made in this configuration or a modification of it. One particular modification is to use electromagnets for the stator and to add additional poles for the magnetic field (also spaced equally). **Figure 5.68** shows a small motor with two stator poles and eight rotating coils (note the way they are wound). The addition of iron increases force and torque. This motor is called a universal motor, and can operate on DC or AC and is perhaps most common in AC-powered hand tools. They can develop a fairly large torque but are very noisy. Typically the stator coils are connected in series with the rotor coils (series universal motors), although parallel connection is also possible. The motor in **Figure 5.68** is a high-speed universal motor as used in a hand tool. It also shows one of the problems common in these motors—damage to the commutator due to sparks developed when brushes (carbon contacts) slide over the commutator in normal operation. These brushes also wear out over time, reducing motor performance.

In many applications, and especially for low-power DC applications, a modification of this basic configuration has the magnetic field produced by a pair (or more) of permanent

FIGURE 5.68 ■ The rotor and stator of a universal motor. Note the method of winding the coils and the damage to the commutators.

magnets and a number of poles produced by windings, as shown in **Figure 5.69**. In this case there are three poles on the rotor and two on the stator (marked as P in **Figure 5.69b**), ensuring that the motor can never get stuck in a zero force situation. The commutator operates as previously, but because there are three coils, one or two coils are energized at a time (depending on rotary position). **Figure 5.69c** shows a similar but somewhat larger motor with seven poles and the same number of contacts on the mechanical commutator. These motors are commonly encountered in tape drives and in toys, as well as in cordless hand tools. They can be reversed by simply reversing the polarity of the source.

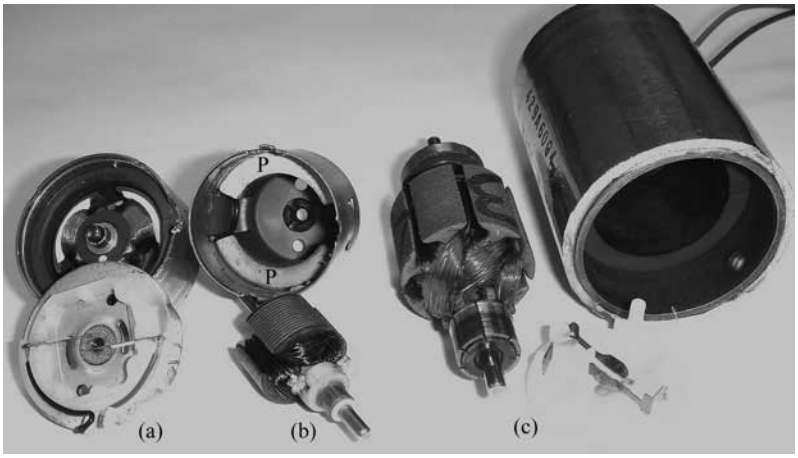


FIGURE 5.69 ■ Three small motors showing the rotors, stators, and commutators. Left and middle: three-pole rotor, two-pole stator. Right: seven-pole rotor, two-pole stator. The sliding "brushes" are seen at the bottom of the figure. In larger motors the brushes are made of carbon or graphite.

EXAMPLE 5.14 Torque in a commutated DC motor

A permanent magnet DC motor based on **Figure 5.67a** has an iron rotor on which the square coil is wound. This guarantees that the magnetic flux density in the gap between the stator and rotor where the coil resides is constant and high. Suppose the magnetic poles produce a magnetic flux density of 0.8 T. The coil is 60 mm long, has a radius of 20 mm, and contains 240 turns.

a. Calculate the maximum torque the motor can generate for a current of 0.1 A in the coils.

b. Suppose now that a second coil is added, as in **Figure 5.67b**. How does the answer in (a) change?

Solution: The fact that the flux density in the coil is constant provides a simple expression for the forces and torque.

a. The torque on the coil depends on the position of the coil, with a maximum when the flux density is parallel to the area of the coil (see **Figure 5.66b** and **Equation (5.65)**) and a minimum when perpendicular. The force on a single loop was calculated above and is available from

Equation (5.26):

$$F = BIL \quad [\text{N}].$$

In the motor discussed here, the length is $L = 0.06$ m, the radius is $r = 0.02$ m, the current is $I = 0.1$ A, and the flux density is $B = 0.8$ T. Since there are $N = 240$ turns in the coil, the total force on the coil is (the force F in **Figure 5.66b**)

$$F = NBIL = 240 \times 0.8 \times 0.1 \times 0.06 = 1.152 \text{ N}.$$

The force is constant regardless of the position of the coil. However, the torque is maximum when the coil is parallel to the field since then the force F is perpendicular to the plane of the coil. Thus the maximum torque is

$$T = 2Fr = 2 \times 1.152 \times 0.02 = 0.046 \text{ N} \cdot \text{m}.$$

b. The torque remains the same as in (a) except that whereas in (a) it varies from maximum to minimum (zero) in one-half turn, in (b) it varies from maximum to minimum in one-quarter turn, producing a smoother torque with the rotation of the motor.

5.9.2.2 Brushless, Electronically Commutated DC Motors (BLDC Motors)

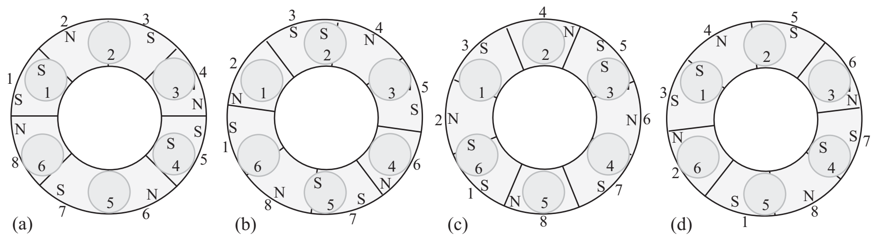
DC motors may be sufficient for simple applications but their control (speed and torque control) is somewhat complex. In addition, the mechanical commutator is electrically noisy (generates sparks and therefore magnetic fields that can interfere with electronic circuits) and wears out with time. For more demanding applications, such as in disk drives, a variation of this motor is used in which the commutation is done electronically. In addition, the physical structure is often different to allow fitting in tight spaces or incorporation with integrated circuits. These motors are often flat (a popular name is flat motors) and often the rotor is a mere disk. An additional important aspect is that now, since commutation is electronic, the coils are stationary and the magnets rotate. These motors can be viewed as a type of stepper motor (we shall talk about stepper motors next), but BLDC motors are typically used for continuous rotation and hence their discussion here.

To understand their operation, consider **Figure 5.70a**. It shows a brushless DC motor with six coils forming the stator. The rotor has been taken out of its bearing and inverted to reveal the stator and the structure of the rotor. The coils are placed directly on a printed circuit board. The rotor, shown on the left, has a ring made of eight separate magnets so that the sides facing the coil (up in this figure) alternate in their magnetic field (the individual magnets can be distinguished by the brighter lines separating them (**Figure 5.70b**)). Note also the three Hall elements placed in the middle of three of the coils—these are used to sense the position of the rotating magnets for the purpose of control of speed and direction of rotation. The operation of the motor relies on two principles. First, the pitch of the stator and rotor are different (six coils but eight magnets). Second, the positions of the magnets are sensed and this information is used to drive the coils, measure the speed, and reverse the sense of rotation. By driving sequential pairs of coils, the device can be made to rotate in one direction or the other. The exact timing for switching the coils is obtained from the three Hall elements. **Figure 5.71** explains the sequence. Suppose that the initial condition is as shown in **Figure 5.71a**. The initial condition of the coils with respect to the magnets is sensed by

(a) (b)

FIGURE 5.70 ■ A flat, electronically commutated DC motor (brushless DC motor or BLDC motor). (a) View of the separated rotor and stator. (b) A closer view of the rotor showing the separation between the individual magnets (lighter strips).

FIGURE 5.71 ■ Operation of a flat brushless DC motor. Circles represent coils, segments represent magnets with alternating polarities (marked).



the Hall elements so that a predictable direction of rotation can be defined. The magnets are shown behind the round coils and polarity is indicated on the periphery to avoid clutter. The coils are shown as grey circles and their polarity (on the side facing the magnets) is indicated in black. Now suppose coils 1 and 4 are driven so that their polarity is as shown in **Figure 5.71a**. That is, their side facing the magnets is *S*. Coil 1 will repel magnet 1 and will attract magnet 2, whereas coil 4 will repel magnet 5 and attract magnet 6. This will rotate the rotor (magnets) to the left until coil 1 is centered with magnet 2 and coil 4 is centered with magnet 6. Next, coils 2 and 5 are driven in the same way (as shown in **Figure 5.71b**). Now coil 2 will repel magnet 3 and attract magnet 4, whereas coil 5 will repel magnet 7 and attract magnet 8. Again, the magnets are forced to rotate left until coil 2 is centered with magnet 4 and coil 5 with magnet 8. This is shown in **Figure 5.71c**. Now the process repeats, but coils 3 and 6 are driven. Again rotation to the left is obtained until the coils and magnets are as in **Figure 5.71d**. This is identical to **Figure 5.71a**, and the process repeats indefinitely. This is said to be a three-phase operation and can be done digitally since all it requires is to ascertain the location of the magnets and drive the opposite coils according to the sequence above. Note that by reversing the coils currents, the north (N) poles are operating against the magnets and rotation is in the opposite direction.

This type of motor is the common choice in most digital devices such as disk drives, CD drives, video recorder heads, tape drives, and many others because it can be controlled very easily and its control is essentially digital. The speed is controlled by timing

FIGURE 5.72 ■ The stator (left) and rotor (right) of an electronically commutated DC motor (from a CD drive). The magnets are placed on the interior rim of the rotor. Note also the three Hall elements used to sense the rotor.

FIGURE 5.73 ■ A sensorless BLDC motor for electrical model aircraft applications. The speed of rotation is up to 70,000 rpm. Note the permanent magnets around the circumference of the stator.

the three phases at will. However, there are many variations in terms of the actual construction, shape, and number of magnets and coils, etc. One additional form is shown in **Figure 5.72**. In this case the magnets are placed on the inner side of the rim of the rotor and the coils are wound on an iron core to increase torque. A three-phase BLDC motor used for electrical model aircraft is shown in **Figure 5.73**. The particular motor shown can rotate at speeds up to 70,000 rpm.

Another development of BLDC motors is to dispense with the Hall elements for sensing and use the coils themselves as sensors for positioning by monitoring the induced emf in those coils that, at any given instant, are not driven. This allows for control of the motor in a sensorless mode. The control is more complicated, but there is no need for separate Hall elements to sense position. The motor in **Figure 5.73** is of this type.

5.9.2.3 AC Motors

In addition to DC motors, there is a large variety of AC motors. The most common of the conventional motors is the induction motor in its many variants. Without going into

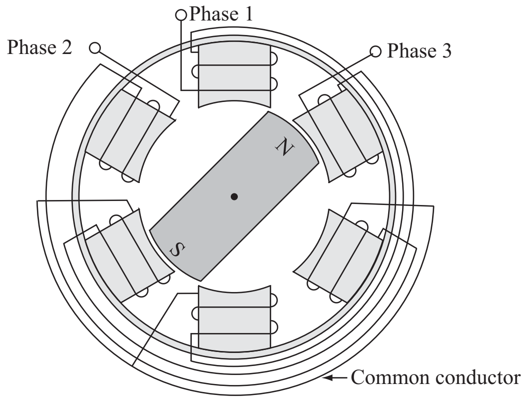


FIGURE 5.74 ■ A motor based on rotation of the magnetic field. In an induction machine, the magnet is replaced by a shorted coil.

all the details of their construction, the induction motor may be understood by first returning to **Figure 5.66**, but now the magnetic flux density is an AC field. In addition, the rotating coil is shorted (no external current connected to it). The AC field and the coil act as a transformer and an AC current is induced in the coil because it is shorted. According to Lentz's law, the current in the coil must produce an opposing field, which then forces the coil to rotate. Since now there is no commutation, continuous rotation is achieved by rotating the field. That is, suppose that an additional field is provided perpendicular to that shown in **Figure 5.66**. This can then be switched on after the loop has rotated one-half turn to keep it going. In practice this is done somewhat differently, by using the phases of the AC power supply to form a rotating field, shown schematically in **Figure 5.74** for a three-phase AC motor (a magnet is shown for the rotor, but a shorted coil acts exactly as a magnet). As the phases of the supply change with time, they generate a rotating magnetic field that drags the rotor with it, affecting the rotation.

Induction machines are very common in appliances since they are quiet, efficient, and most importantly, rotate at constant speeds that depend only on the frequency of the field and the number of poles. They are also used in control devices where constant speed is important. Control of induction motors, other than on and off, is much more involved than for DC motors, especially when variable speed is desired. A small induction motor is shown in **Figure 5.75**.

Of course, there are other types of AC motors with particular performance characteristics.

5.9.2.4 Stepper Motors

In actuation, the control of a motor for, say, exact and repeatable positioning requires, in addition to a continuous motor, some means of feedback, counting rotations, sensing position, etc. Motors that incorporate these means are called **servomotors** and they are commonly used in various control systems. In some applications servomotors have been replaced by stepper motors. Stepper motors are incremental rotation or linear motion motors. For this reason they are often viewed as "digital" motors, in the sense that each increment is fixed in size and increments are generated by a train of pulses. To understand their operation, consider first the configuration in **Figure 5.76a**. This is a two-phase stepper motor and uses a permanent magnet as the rotor, allowing a simple description of the operation. The rotor can be made to rotate in steps by properly driving the two coils, which in turn define the magnetic poles of the stator. To see the stepping

FIGURE 5.75 ■ A small induction motor.

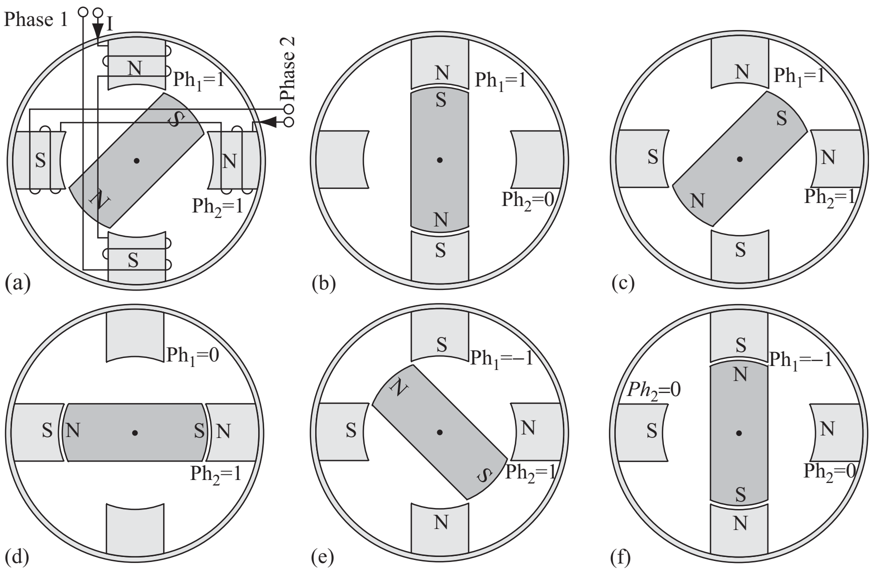


FIGURE 5.76 ■ (a) A schematic of two-phase stepping motor. The currents shown correspond to both phases being driven ($Ph_1 = 1$, $Ph_2 = 1$). (b–f) The sequence of half-stepping for the two phase motor in (a).

sequence, consider **Figure 5.76b**. By driving the two vertical coils, the magnet is held aligned vertically with coil 1. Now, if both coils are driven as in **Figure 5.76c**, the rotor will be at rest at 45°, rotating to the right. This is called a half-step and is the minimum rotation or step possible in this stepper motor. If now the vertical coil is deenergized, but the horizontal coil is kept energized, the magnet rotates an additional quarter turn to the position in **Figure 5.76d**. In the next step the current in the vertical coil is negative, in the horizontal coil it is positive, and the situation in **Figure 5.76e** is obtained. Finally, by reversing the vertical coil current and setting the horizontal coil to zero (no current), a full rotation has been completed (**Figure 5.76f**). This simple motor steps at 45° and requires eight steps to rotate one full turn. To rotate in the opposite direction the

TABLE 5.6 ■ The sequence required to turn the motor in **Figure 5.76** one full rotation (eight steps) clockwise[a]

Step	S_1	S_2
1	1	1
2	0	1
3	−1	1
4	−1	0
5	−1	−1
6	0	−1
7	1	−1
8	1	0

0 – no current, 1 current in one direction, –1 current in opposite direction.
[a]Reversing the sequence turns it counterclockwise. Shaded rows show the full step sequence.

sequence must be reversed, as shown in **Table 5.7**. The sequence above indicates the following:

1. The size of the step (number of steps) depends on the number of coils and, as we shall see later, the number of poles in the rotor.
2. Full stepping (90° in this case) can be accomplished by using only one of the stator coils (single phase) at each step.
3. More coils and more poles in the rotor will produce smaller steps.
4. The number of poles in the rotor and in the stator must be different (fewer poles in the rotor).
5. The magnetic field in the rotor can be generated by permanent magnets or by coils. We shall see that neither is really necessary. A stepper motor can be made with an iron rotor alone.

From the previous discussion it is clear that whereas the structure in **Figure 5.76** is capable of half-stepping, it can also be used for full steps (i.e., at increments of 90°). This is done by skipping steps 1, 3, 5, and 7 in **Table 5.6**. The sequence is shown in the shaded rows. That is, the same stepping motor can be used to move faster or slower through the sequence.

Consider again **Figure 5.76**, but suppose that the permanent magnet in the rotor is replaced with a piece of iron (nonmagnetized). The operation indicated above is still valid since the magnetic field produced by the stator coils will magnetize the iron (i.e., an electromagnet will attract a piece of iron). This simplifies matters considerably since now the rotor is much simpler to make. This type of stepper motor is called a **variable reluctance stepping motor** and is a common way of producing stepper motors. The principle of the motor is shown in **Figure 5.77a**. To operate it, the coils marked as 2 are first energized. This moves the rotor one step counterclockwise. Then coils marked as 3 are energized, moving one step to the left, and so on. Rotation in the opposite direction is obtained by inverting the sequence (driving coil number 3 first, then 2, and so on).

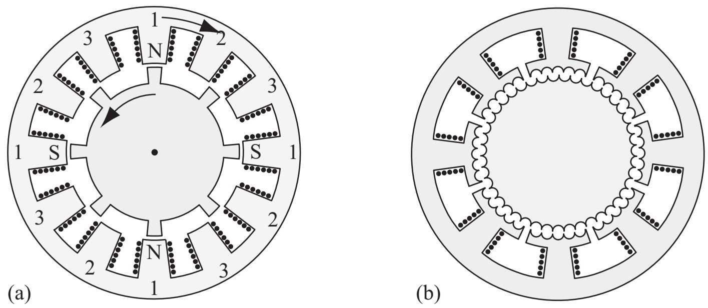


FIGURE 5.77 ■ (a) A practical stepper motor with 12 poles in the stator and 8 poles in the rotor. This motor is a three-phase motor. (b) An 8-pole (40 teeth) stator and 50-teeth rotor variable reluctance stepping motor.

Assuming there are n_s stator poles and n_r rotor poles (teeth in this case). The stator and rotor pitches are defined as

$$\theta_s = \frac{360^\circ}{n_s},\ \theta_r = \frac{360^\circ}{n_r}. \tag{5.66}$$

The step of the stepper motor is then

$$\Delta\theta = |\theta_r - \theta_s|. \tag{5.67}$$

In the example in **Figure 5.77a** there are 12 poles in the stator and 8 in the rotor. Thus the full-stepping angle can be calculated as

$$\theta_s = \frac{360^\circ}{12} = 30^\circ,\ \theta_r = \frac{360^\circ}{8} = 45^\circ,\ \Delta\theta = |45^\circ - 30^\circ| = 15^\circ. \tag{5.68}$$

Half-stepping is possible by proper driving following the principles outlined above. For example, by driving first coil 1 we obtain the situation shown in the figure. If coils 1 and 2 are driven at the same time, the rotor will move counterclockwise one-half step (see **Example 5.16**). This stepper motor is capable of full steps of 15°. The motor is a three-phase stepper motor since the sequence to drive it in either direction repeats after three steps. Note that here the number of poles in the stator is larger than in the rotor. The opposite is just as valid and is often used in conjunction with variable reluctance stepper motors.

To simplify construction, the rotor is made of n_r teeth, as above, and the stator is made of a fixed number of poles, say eight, and each pole is toothed, as shown in **Figure 5.77b**. Note that in this case there are more teeth in the rotor (50) than in the stator (40). This produces a step of 1.8° (360/40−360/50). The motor in this figure is a four-phase motor (it requires a sequence of four pulses that repeat indefinitely).

EXAMPLE 5.15 A 200 step/revolution motor

One the most common stepper motors is the 200 steps/revolution motor. It is typically made of eight poles in the stator, each split into five teeth (a total of 40 teeth). The rotor has 50 teeth, as in **Figure 5.77b**. The full-step angle is calculated as follows:

In the stator,

$$\theta_s = \frac{360^\circ}{40} = 9^\circ.$$

In the rotor,

$$\theta_r = \frac{360^\circ}{50} = 7.2^\circ.$$

Thus the full step is

$$\Delta\theta = 9^\circ - 7.2^\circ = 1.8^\circ$$

and the number of steps per revolution is $360^\circ/1.8^\circ = 200$.

Note: By splitting the rotor in two and shifting half of the rotor by 1/tooth, the same motor is capable of half-stepping (0.9°/step or 400 steps/revolution). It is also possible to half-step by proper driving of the coils without the need to split the rotor, as discussed above and in the following example.

EXAMPLE 5.16 Half-stepping of a variable reluctance stepper motor

Consider the variable reluctance stepper motor shown in **Figure 5.77a**. Show the driving sequence to produce half-step motion in the clockwise direction.

Solution: We start from the position shown, that is, the first step in the sequence is to drive coil 1 as shown. To move clockwise, we need to attract the tooth to the left of the aligned tooth, meaning we must drive coils 1 and 3. The condition now is shown in **Figure 5.78a**. Next we drive coil 3 to obtain the configuration in **Figure 5.78b**. Now coils 3 and 2 are driven at the same time (**Figure 5.78c**), followed by coil 2, then coils 1 and 2 together. The next step is to drive coil 1 alone to get to the starting step in **Figure 5.77a**. Thus the sequence is (1), $(1+3)$, (3), $(3+2)$, (2), $(2+1)$.

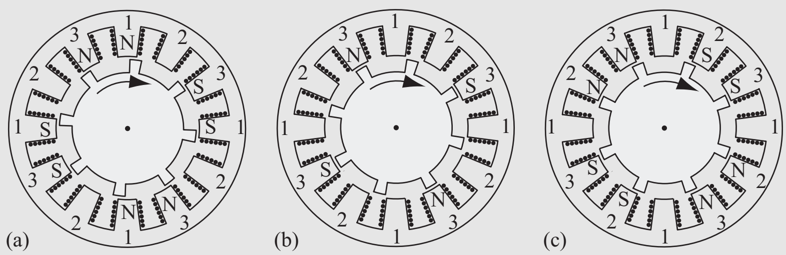


FIGURE 5.78 ■ Sequence necessary for half-stepping of the variable reluctance stepper motor in **Figure 5.77a**. Starting with the position in **Figure 5.77a,** (a–c) show the first three half-steps and the necessary coil driving sequence.

FIGURE 5.79 ■ A double-stack stepper motor with permanent magnet rotor (1.8° steps).

In general, variable reluctance stepping motors are simpler and less expensive to produce. However, when not powered, their rotor is free to move and hence they cannot hold their position. Permanent magnet rotors in stepper motors have some holding power and will maintain their position under power-off conditions.

In the description above, the stepper motor had a single rotor and a single stator. In an attempt to make motors with increasingly finer pitch (smaller steps), multiple rotors on a single shaft are used. These are called multiple stack stepper motors and they allow much finer steps by keeping one of the set of poles (on the rotor or on the stator) equal but changing the pitch on the opposite set of poles. Since now the pitch varies between stacks rather than within one stack, finer pitches are possible. The disadvantage is that the driving sequence is more complicated than in a single rotor motor. Usually the stator and each rotor have the same number of teeth but the two rotors are shifted one-half tooth apart (for a two-stack rotor). An example of an eight-pole (stator), double-stack motor is shown in **Figure 5.79**. This motor has 50 teeth on each rotor and 40 teeth on the stator. The rotors are magnetized and the motor shown has a 1.8° step but is capable of 0.9° by proper driving. We shall not pursue these motors further since they are not fundamentally different than single-stack motors.

Stepper motors come in all sizes, from tiny to very large, and are currently the motors of choice for accurate positioning and driving. They are, however, by themselves, more expensive and lower power than other motors such as DC motors. The extra cost is usually justified by their simple control and accuracy and by the fact that they can be driven from digital controllers with little more than a transistor or metal oxide semiconductor field-effect transistor (MOSFET) per phase to supply the current needed.

One can find stepper motors in industrial controls as easily as in consumer products such as printers, scanners, and cameras. In these applications, the ability of the motor to step through a predictable sequence with accurate, repeatable steps is used for fast positioning. The motors have typically low inertia, allowing them to respond quickly in both directions. In this respect they are fully capable of incorporation in fast systems while still maintaining their direct drive capabilities. Two small stepping motors are shown in **Figure 5.80**.

FIGURE 5.80 ■ Two small stepper motors. Left: A motor from a disk drive. Right: A paper advance motor from an ink jet printer.

5.9.2.5 Linear Motors

Conventional motors are naturally suited for rotary motion. There is, however, a need for linear actuation that cannot be met by rotary motors directly. In such cases, the rotary motion can be converted to linear motion through the use of cams, screw drives, belts, and so on. Another possibility, one that is gaining popularity, is the use of linear motors. We have in fact discussed two methods of linear motion in previous sections (voice coil actuators and the inchworm magnetostrictive motor).

A linear motor, either incorporating continuous motion or stepper motion, can be viewed as a rotary motor that has been cut and flattened so that the rotor can now slide linearly over the stator. This is shown schematically in **Figure 5.81**. Note that the slider or translator (equivalent to the rotor) may have as many poles as we wish—four are shown for clarity. Starting from the initial condition in **Figure 5.81a**, the sliding poles are driven as shown and are therefore attracted to the right. As they pass past the stator poles, they are commutated and the polarities change as in **Figure 5.81b**, again forcing motion to the right. This is merely a commutated DC machine. Motion to the left requires the opposite sequence. Based on this description, any of the motors above, including induction motors, can be built as linear motors. The stator may be very long (such as in linear drives for trains—in which the stator equals the length of the rail) or may be fairly short, depending on the application.

A variable reluctance linear stepping motor may be built as shown in **Figure 5.82**. This motor is equivalent to the rotary motor in **Figure 5.77a**. However, now the pitch is measured in units of length (so many millimeter per step). In this sequence we assume that the stator poles are driven and that the rotor is a mere toothed iron piece (variable reluctance motor). The sequence is as follows: Starting with the configuration in

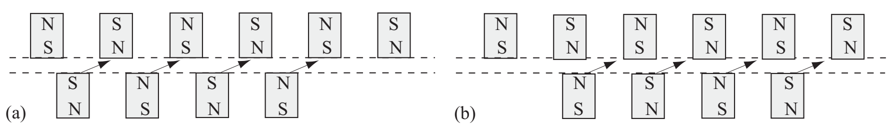


FIGURE 5.81 ■ The principle of linear motors. (a) The translator (bottom part) moves to the right until it is centered under the stator's poles. (b) The polarities of the translator are commutated and a new step can take place.

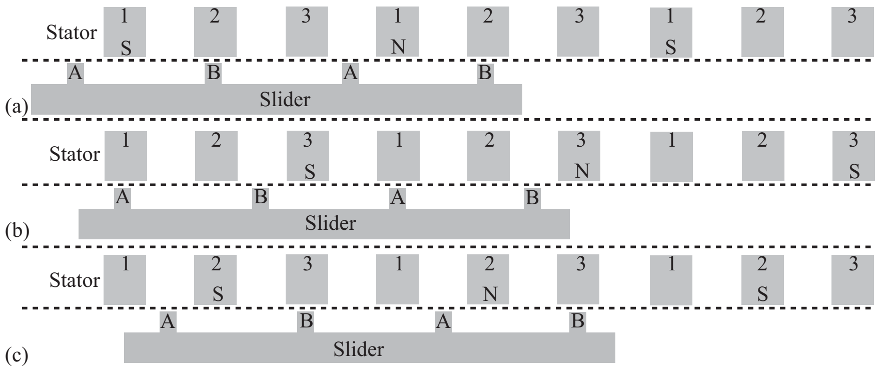


FIGURE 5.82 ■ Operation of a three-phase linear stepper motor with driven stator. Sequence is 1-3-2.

Figure 5.82a, poles marked as 1 are driven alternately as N and S, as shown. The slider moves to the right until teeth A are aligned with poles 1. This is shown in **Figure 5.82b**. Now poles 3 are driven as previously and the slider again moves to the right until teeth B are aligned with poles 3 (**Figure 5.82c**). Finally, poles 2 are driven, and the cycle completes and the relation of the slider and the stator is now as at the beginning of the sequence. The same can be accomplished with permanent magnet poles in the rotor. From **Figure 5.82** it should be noted that the pitch of the stator and slider are different. For every four poles in the stator there are three teeth in the slider. Thus each step is equal to half the pitch of the stator (i.e., in each step a tooth moves either from the middle between two poles to the center of the next pole, or vice versa). Of course, by changing the number of teeth, one can change this pitch. In the motor described here, the sequence is 1-3-2 for motion to the right. Moving in the opposite direction is accomplished by changing the sequence to (3-1-2).

In many linear stepping motors it is more practical to drive the slider rather than the stator since the stator may be very long, whereas the slider is usually short. However, the principle is the same. A variable reluctance linear stepping motor (which includes permanent magnet poles) is shown in **Figure 5.83** with the slider off the stator and inverted to reveal its poles (four poles with six teeth on each pole). These are separated 1 mm apart. The teeth on the stator are slightly smaller, allowing for the fine step motion of this device.

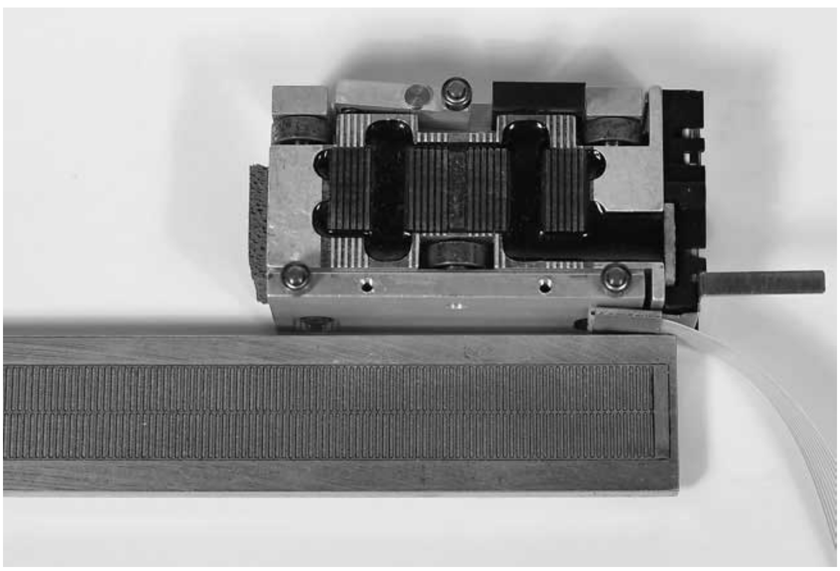


FIGURE 5.83 ■ A disassembled linear stepper motor with drive translator showing the poles of the translator and those of the stator.

There are many other issues involved in the use of motors as actuators, some mechanical, some electrical. Some aspects of their use, including starting methods for AC machines, inverters, power supplies, protection methods, and the like are outside the scope of this text.

5.9.3 Magnetic Solenoid Actuators and Magnetic Valves

Magnetic solenoid actuators are electromagnets designed to affect linear motion by exploiting the force an electromagnet can generate on a ferromagnetic material. To understand the operation, consider the configuration in **Figure 5.84a**. A coil generates a magnetic field everywhere, including in the gap between the fixed and movable iron pieces. We shall call the movable piece a plunger. In a closed magnetic path, such as the solenoid in **Figure 5.84b**, the magnetic flux density in the air gap between the plunger and the fixed iron piece is (approximately)

$$B = \frac{\mu_0 NI}{L} \quad [\mathrm{T}], \tag{5.69}$$

where N is the number of turns in the coil, I is the current in the coil, and L is the length of the gap (see **Figure 5.84a**). This is approximate because it neglects the effect of the field outside the air gap between the plunger and the fixed piece and does not apply when L approaches zero or when L is large. Nevertheless, it provides a good approximation for the magnetic flux density in many cases, especially in the case shown in **Figure 5.84b** in which the magnetic field closes through iron because of its high permeability. The method used here is known as a "virtual displacement method" and is similar to the method used to obtain the electrostatic force in **Equation (5.14)**. The force exerted on the plunger is given as

$$F = \frac{B^2 S}{\mu_0} = \frac{\mu_0 N^2 I^2 S}{L^2} \quad [\mathrm{N}], \tag{5.70}$$

where B is the magnetic flux density in the gap generated by the coil (perpendicular to the surface of the iron pieces), S is the cross-sectional area of the plunger, and μ_0 is the permeability of free space (air) in the gap.

The force on the plunger tends to close the gap and this motion is the linear motion generated by the magnetic valve actuator. As the plunger closes the gap, the force increases because L decreases. The construction shown in **Figure 5.84b** is more practical since it generates an axial field in the plunger and closes the external field so that

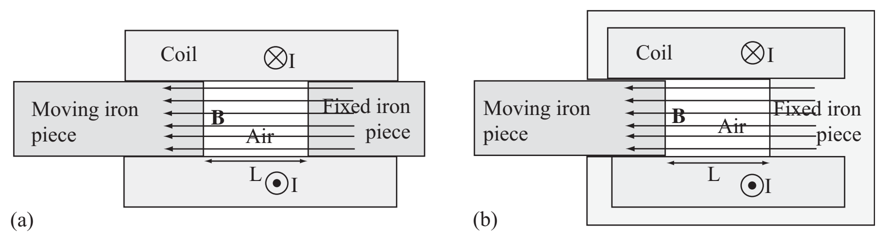


FIGURE 5.84 ■ The solenoid actuator. (a) The coil generates a magnetic field in the gap between the fixed and movable pieces. (b) A more practical construction that ensures the flux is closed and increases the magnetic flux density, and hence the force.

the total magnetic field available at the plunger is larger. In this form, the device is used as a simple go/no go actuator. That is, when energized, the gap is closed, and when deenergized it is open. This type of device is often used for electrical release of latches on doors, as a means of opening/closing fluid or gas valves, or to engage a mechanism. Examples of small linear solenoid actuators are shown in **Figure 5.85**. A modification of the linear plunger is the rotary or angular solenoid actuator, an example of which is shown in **Figure 5.86**. In this example, the rotor can move one-half turn in either direction. The rotor, which is equivalent to the plunger in the linear case, is made of a permanent magnet to increase the force.

The basic solenoid actuator is often used as the moving mechanism in valves. A basic configuration is shown in **Figure 5.87**. These valves are quite common in

FIGURE 5.85 ■ Two linear solenoid actuators. The plungers are shown in the middle of the devices and are attached to whatever is actuated by them.

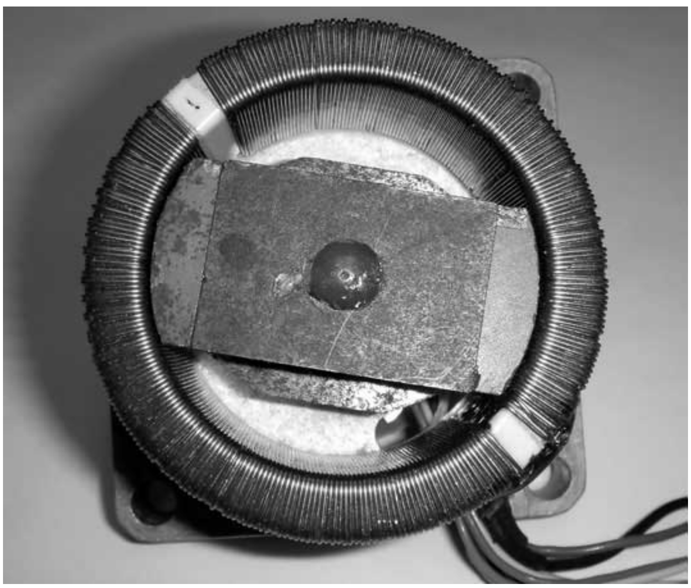


FIGURE 5.86 ■ An angular solenoid actuator. Here the rotor (equivalent to the plunger) is a permanent magnet.

control of both fluids and gases and exist in a variety of sizes, constructions, and power levels. They can be found not only in industrial processes, but also in consumer appliances such as washing machines, dishwashers, and refrigerators, as well as in cars and a variety of other products. The actuating rod (plunger) in this case acts against a spring and by properly driving the current through the solenoid its motion can be controlled as to speed and force exerted. Similar constructions can operate and control almost anything that requires linear (or rotational) motion. However, the travel of the actuating rods is relatively small, on the order of 10–20 mm, and often much less.

A magnetic valve designed for fluid flow control is shown in **Figure 5.88**. **Figure 5.89** shows a smaller valve used for control of air flow. The solenoid is about 18 mm in diameter, 25 mm long, and operates at 1.4 V and 300 mA.

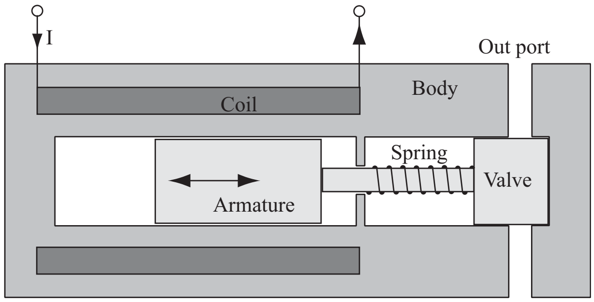
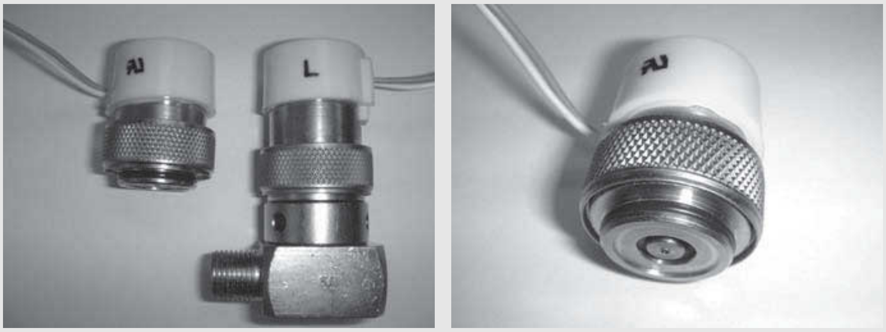


FIGURE 5.87 ■ The principle of a valve solenoid actuator, showing the coil and return spring. In this case the valve closes or opens an orifice.

FIGURE 5.88 ■ An electric valve for fluid flow control operated by a magnetic coil (28 V DC or 110 V AC).

EXAMPLE 5.17 **Force produced by a linear solenoid actuator**

A solenoid actuator has a cylindrical plunger of diameter 18 mm and a total travel (L in **Figure 5.84b**) of 10 mm. The coil contains 2000 turns and is fed with a constant current of 500 mA. Calculate the initial force the solenoid can exert (i.e., when the gap is 10 mm) and after the plunger has traveled 5 mm.



(a) (b)

FIGURE 5.89 ■ A valve used to control air flow to a piston. The solenoid operates at 1.4 V and 300 mA. (a) Solenoid and valve. (b) Detail of the solenoid. The solenoid is 18 mm in diameter and 25 mm long.

Solution: The initial force is the most important parameter since it is this force that must act to affect action (such as opening a lock). It can be calculated directly from **Equation (5.70)** with $N = 2000$, $I = 0.5$ A, $L = 0.01$ m, and $\mu_0 = 4\pi \times 10^{-7}$ H/m. This produces a force of

$$F = \frac{\mu_0 N^2 I^2 S}{L^2} = \frac{4\pi \times 10^{-7} \times 2000^2 \times 0.5^2 \times (\pi \times 0.009^2)}{0.01^2} = 3.198 \text{ N}.$$

After a travel of 5 mm, $L = 0.005$ and the force is

$$F = \frac{\mu_0 N^2 I^2 S}{L^2} = \frac{4\pi \times 10^{-7} \times 2000^2 \times 0.5^2 \times (\pi \times 0.009^2)}{0.005^2} = 12.791 \text{ N}.$$

Not surprisingly, this force is four times larger since L is twice as small.

Notes: The forces these solenoids produce are not large, but are sufficient to open a valve, unlock a door, or pull a mechanical lever to release a device. On the other hand, they typically dissipate relatively large amounts of power in the coil and hence tend to be used intermittently. However, there are solenoids that can be turned on continuously. Note also that the calculation performed here is only valid for $L > 0$ because of the assumptions used in the development of **Equation (5.70)**. A more exact calculation can be done by taking into account the effect of the iron path and the magnetic properties of iron.

5.10 | VOLTAGE AND CURRENT SENSORS

In most cases voltage and current are measured as output of sensors or applied to actuators. But the sensing of voltage and current are important in themselves, as these are often used to affect other conditions. For example, the control and regulation of the output of a power supply requires that the output voltage and current be monitored. To maintain a constant voltage on a power line or to regulate the voltage in a car requires similar sensing of voltage and current. Fuses and circuit breakers are devices that sense the current in electric circuits and disconnect the power to the circuit when the current exceeds a preset value. Other devices protect circuits from overvoltages.

There are many mechanisms for sensing current and voltage. The most common methods are resistive and inductive, but the Hall element can be used successfully (see **Example 5.20**), as can capacitive methods. Some of the principles for DC and AC voltage and current sensing are discussed next.

5.10.1 Voltage Sensing

The potentiometer is a variable voltage divider, shown schematically in **Figure 5.90**. Although the purpose of using a potentiometer may vary, in all cases an input voltage V_{in} is divided to produce an output voltage V_{out}, which may be viewed as a "sampling" of the voltage V_{in}. That is,

$$V_{\text{out}} = \frac{V_{\text{in}}}{R} R_o \quad [\text{V}]. \tag{5.71}$$

Potentiometers come in a variety of physical implementations. The potentiometer may be a rotary or a linear device and its sampling may be linear or nonlinear, of which logarithmic potentiometers are the most common. In a rotary potentiometer, a resistance is built on a circular path and the slider rotates on a shaft to sample part of the resistance to produce the output. In a linear potentiometer, the resistance is a straight strip and the slider moves from side to side. A logarithmic-scale potentiometer's resistance is nonlinear, varying on a logarithmic scale (see **Example 2.15** for a discussion of the logarithmic potentiometer) and may be linear or rotary. Similarly, a linear-scale potentiometer may be linear or rotary. There are also potentiometers with multiturn capabilities and potentiometers without a shaft (often called trimmers) intended for one-time or occasional adjustment, usually using a screwdriver. Electronic potentiometers are devices that produce the same effect, that of sampling part of a voltage by electronic means. **Figure 5.91** shows a number of potentiometers of various types and sizes. The potentiometer allows one to sense the voltage V_{in} in **Figure 5.90**, which may be high, through use of the voltage V_{out}, which may be adjusted to a convenient level. Sometimes, when V_{in} is very high, the use of a voltage divider is a must. The potentiometer is equally suitable for DC and AC voltage sensing.

The common transformer is another means of sensing voltage, but now the voltage must be AC. The transformer was discussed in **Section 5.4.1** and is shown schematically in **Figure 5.92**. The output voltage relates to the input voltage through the turn ratio, as shown in **Equation (5.31)**. With **Figure 5.92a** as a reference, the output voltage is

$$V_{\text{out}} = \frac{V_{\text{in}}}{N_1} N_2 \quad [\text{V}]. \tag{5.72}$$

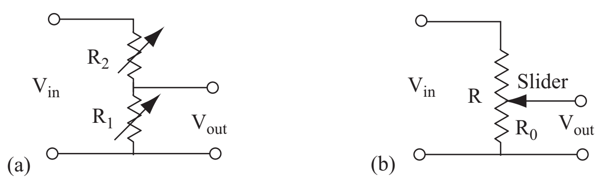


FIGURE 5.90 ■ The potentiometer as a resistive voltage divider. (a) Two variable resistors can produce any output voltage between zero and V_{in}. (b) The potentiometer does this by varying the ratio between the two resistors while their sum remains constant.

FIGURE 5.91 ■ A number of potentiometers of various types and sizes.

FIGURE 5.92 ■ The transformer. (a) Common isolating voltage transformer. (b) The autotransformer. Both serve as voltage sensors.

Unlike the potentiometer, the transformer also isolates the sampled voltage from the input voltage, a property that is important when mixing high and low voltages, and in particular where contact with the input voltage is to be avoided, usually for safety reasons. Although variable transformers exist (**Figure 5.92b**), they are not common, and unlike the standard transformer, which has a constant turn ratio, the variable transformer has a variable turn ratio. In fact, it is a type of potentiometer. In spite of its apparent usefulness, variable transformers are not often used because they are bulky, expensive, and most of the designs do not isolate input and output.

Voltage can also be sensed capacitively in what can be called a capacitive voltage divider, shown in **Figure 5.93a**. The output voltage is given as

$$V_{out} = \frac{C_2}{C_1 + C_2} V_{in} \quad [\mathrm{V}]. \tag{5.73}$$

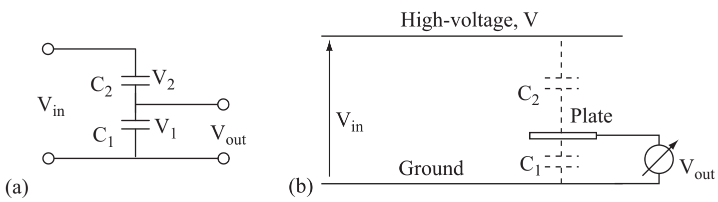


FIGURE 5.93 ■ The capacitive divider as a voltage sensor. (a) The principle. (b) An example of sensing the voltage of an overhead power line.

The capacitive method is particularly useful for measurement (sampling) of high voltages where direct measurements are not possible. It works equally well with DC or AC sources. For example, one can envision sensing the voltage of a high-voltage line in a high-voltage device as in **Figure 5.93b**. A conducting wire or a small plate at some height above the ground establishes the capacitance C_1. The capacitance between the high-voltage line and the plate establishes C_2. Once these two capacitances are measured (or are known), the voltage on the plate can be calibrated to monitor the voltage on the line. After that the line voltage variations can be sensed on the plate.

EXAMPLE 5.18 Monitoring of voltage in a high-voltage power supply

There are many installations that make use of high-voltage sources for particular applications. For example, sandpaper is often produced by attracting the abrasive particles using voltages in excess of 100 kV to the paper after applying a layer of glue. Consider a device of this type shown in **Figure 5.94**. The high-voltage supply is applied between the top and bottom surfaces, which are 2 m long, 1.2 m wide, and separated 10 cm apart, forming a capacitor. The abrasive particles are placed on the bottom surface and attracted to the top, sticking to the paper. A small plate of area S is placed at a small distance from the lower conducting surface. The potential difference between the plate and ground is connected to a microprocessor to monitor the high voltage across the plate. If the high voltage can vary between 0 and 200 kV, what must be the distance of the small plate from the bottom surface if the microprocessor operates at 5 V?

Solution: The large plates are fairly close to each other, forming a parallel plate capacitor. Therefore the electric field intensity between the plates is uniform and both the capacitor C_1 and C_2 may be considered as parallel plate capacitors in spite of the fact that the area of the small plate may not be large. Assuming the area is S, we have (as an approximation)

$$C_1 = \frac{\varepsilon_0 S}{d_1}, C_2 = \frac{\varepsilon_0 S}{d_2} \quad [\text{F}].$$

The output must not exceed 5 V at an input voltage of 200 kV. Therefore

$$V_{out} = \frac{C_2}{C_1 + C_2} V_{in} = 5\text{ V} = \frac{\dfrac{\varepsilon_0 S}{d_2}}{\dfrac{\varepsilon_0 S}{d_1} + \dfrac{\varepsilon_0 S}{d_2}} \times 200 \times 10^3 \rightarrow V_{out} = \frac{d_1}{d_1 + d_2} \times 200 \times 10^3 = 5\text{ V}.$$

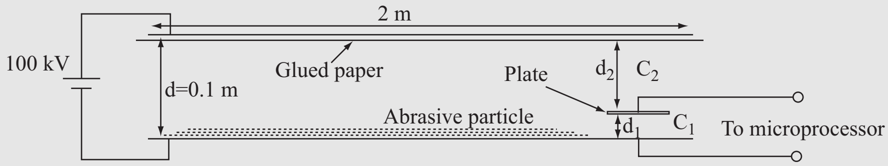


FIGURE 5.94 ■ A small capacitor used as a voltage sensor in a sandpaper production machine.

Since $d_1 + d_2 = d = 100$ mm, we can write

$$V_1 = \frac{200 \times 10^3}{100} \times d_1 = 5 \rightarrow d_1 = \frac{500}{200,000} = 0.0025 \text{ mm}.$$

Note that the area of the plate is immaterial. Also, the plate would have to be protected from dust. But aside from these minor difficulties, the method is viable.

5.10.2 Current Sensing

Most current sensors are in fact voltage sensors, or current to voltage converters. Here also there are a number of methods that can be used. In its simplest form, a resistor connected in series with the current to be sensed provides a voltage proportional to the current, as in **Figure 5.95**. This simple method is often used to sense current in power supplies, in electric machines, and in converters where the sensed current is used to control current or power. The sensing resistor has to be small so as not to affect the current and the voltage of the device, and is typically a fraction of an ohm, but depending on the current, it must be sufficiently large to produce a voltage drop on the order of at least 10–100 mV. This voltage can be amplified to produce the necessary control voltage.

A second method often used with AC currents is the so-called current transformer, shown in **Figure 5.96a**. In fact, it is a regular transformer with the current being sensed being the single-turn primary and with N_2 turns in the secondary. The current I produces a voltage across the primary and that voltage produces a voltage N_2 times larger in the secondary (see **Figure 5.96b**). This voltage is measured and is then an indication of the current in the conductor. There are two basic types of current transformers. One is a solid core transformer, as in **Figure 5.96a**, that requires the sensed current to pass through the

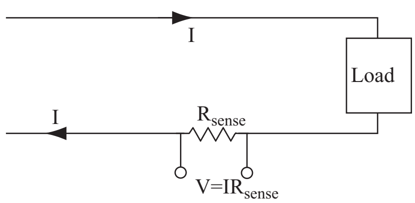


FIGURE 5.95 ■ A resistor as a current sensor. Measuring the voltage on a small, known resistor indicates the current in the load.

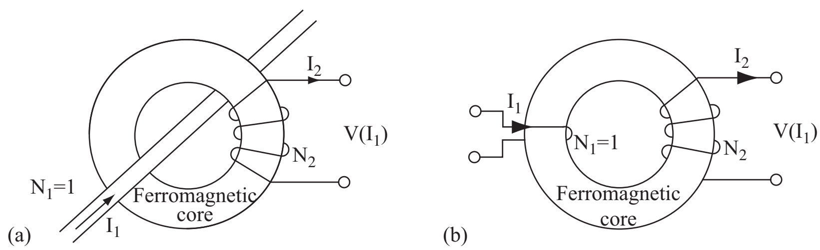


FIGURE 5.96 ■ (a) A current transformer as a current sensor. (b) The equivalent circuit with the current-carrying conductor shown as a single loop.

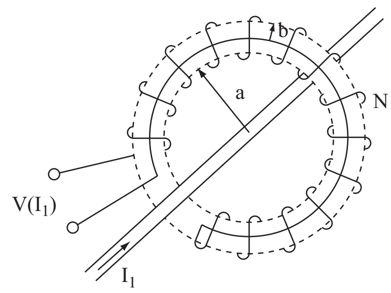


FIGURE 5.97 ■ The Rogowski coil as a current sensor. Since it is not a closed coil, it can fit around conductors more easily than the current transformer in **Figure 5.96**.

core. To facilitate use, some current transformers have hinged cores that can be opened in a manner similar to a clip and closed over the conductor through which the current is measured. The second type uses no core and is based on the Rogowski coil, shown in **Figure 5.97**. The coil is wound uniformly around a round form that is then removed and the end of the wire threaded through the coil itself so that both ends of the coil are available at one end of the coil. This means that the Rogowski coil can be placed over the conductor, facilitating measurements. The coil itself can be potted for physical protection and to maintain its form. The sensor is based on the fact that a current-carrying wire produces a magnetic flux density **B**, given in **Equation (5.17)**. The Rogowski coil has no core, therefore the permeability is μ_0 (that of air or the potting material, usually a plastic). If the coil has an average radius a, the flux density at the center of the coil is

$$B = \mu_0 \frac{I}{2\pi a} \quad [\mathrm{T}]. \tag{5.74}$$

The measured quantity is the emf in the coil. Assuming the turns of the coil are of radius b and there are N turns in the coil, the emf is calculated from **Equation (5.60)**, so we first need to calculate the flux:

$$\Phi = \int_S B ds \approx BS = B\pi b^2 = \frac{\mu_0 I b^2}{2a} \quad [\mathrm{Wb}]. \tag{5.75}$$

If the current is time dependent, and we will assume it to be of the form $I(t) = I_0 \sin(\omega t)$, where $\omega = 2\pi f$ and f is the frequency, we get from **Equation (5.60),**

$$emf = N\frac{d\Phi}{dt} = N\frac{\mu_0 I_0 b^2}{2a}\omega\cos\omega t \quad [\mathrm{V}] \tag{5.76}$$

or

$$emf = \left(N\frac{\mu_0 b^2}{2a}\omega\right) I_0 \cos\omega t \quad [\mathrm{V}]. \tag{5.77}$$

This provides a voltage that is linear with respect to current and can be sufficiently large to be measured directly or after amplification. Note also that the higher the frequency, the higher the output voltage.

If instead of the Rogowski coil one uses the ferromagnetic core coil shown in **Figure 5.96a**, the permeability of free space, μ_0, is replaced with the permeability of the ferromagnetic core, μ, in all relations above. This produces a larger emf for the same number of turns (or fewer turns may be used to obtain the same emf) because the permeability of ferromagnetic materials is much higher. In most applications the ferromagnetic core is in the form of a toroid with an average radius a and cross-sectional radius b and the turns wound uniformly around the core. The only disadvantage of this arrangement is that the current-carrying wire whose current is measured must be threaded through the core, unless the core is hinged so it can be opened and closed around the wire as is done in hand-held clamp current probes.

EXAMPLE 5.19 Current sensor for house power monitoring

A current sensor based on the Rogowski coil is needed to sense the current entering a home. With a maximum expected current of 200 A (root mean square [RMS]), it is desired to use a sensor that produces a maximum voltage of magnitude 200 mV so that it can be connected directly to a digital voltmeter and the current read directly on the 0–200 mV scale. Design a Rogowski coil that will accomplish this. The diameter of the current-carrying conductor is 8 mm and the AC in the grid is sinusoidal at 60 Hz.

Solution: The output of the Rogowski coil in **Equation (5.77)** indicates that we can control three parameters: the average radius of the coil, a (but a must be larger than 4 mm so it can fit over the conductor), the radius of the turns b, and the number of turns. Since a is in the denominator it should be as large as practical. We will arbitrarily use a 5 cm diameter for the coil so that $a = 0.025$ m. Then we solve for b^2N and then decide on b and N so we obtain reasonable results. The peak current, by definition, is $I_0 = 200\sqrt{2}/2 = 282.84$ A. However, we will use the RMS value since the voltage is needed as RMS value.

The maximum emf is

$$emf = N\frac{\mu_0 I_0 b^2 \omega}{2a} = \frac{4\pi \times 10^{-7} \times 200 \times 2\pi \times 60}{2 \times 0.025} b^2 N = 0.2 \rightarrow b^2 N = \frac{0.02}{0.192\pi^2} = 0.1055.$$

That is, the product b^2N must be 0.1055. Since b cannot be very large we will use a 10 mm diameter for the coil ($b = 0.005$ m). This gives

$$N = \frac{0.1055}{b^2} = \frac{0.1055}{0.005^2} = 4222 \text{ turns}.$$

This is a large number of turns, but it is not impractical. Since magnet wire with diameters less than 0.05 mm is available, the coil will require about two layers of closely wound wires. A thicker wire, say 0.1 mm in diameter, would require about four layers of closely wound turns.

The parameters used here can be changed. Using a larger diameter coil would require a greater number of turns, whereas a larger turn diameter would require fewer turns. Note also that if we were to use a high-permeability core, that would reduce the number of turns by a factor equal to the relative permeability of the core, but then the coil would be closed and would not be a Rogowski coil.

Equation (5.17) (or **Equation (5.74)**) shows that the magnetic flux density B produced by a long wire carrying a current I is directly proportional to current and inversely proportional to the distance r from the wire. A current sensor can be devised by measuring the magnetic flux density produced by the current:

$$I = \left(\frac{2\pi r}{\mu}\right) B \quad [\mathrm{A}], \tag{5.78}$$

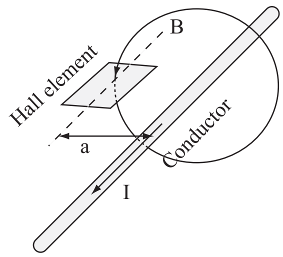


FIGURE 5.98 ■ Principle of a current sensor utilizing a Hall element. The Hall element is embedded in a plastic ring (not shown) through which the conductor passes.

where r is the distance from the conductor at which the magnetic flux density is measured and μ is the permeability at that location. The magnetic flux density may be measured using a small coil, but more often it is measured using a Hall element placed with its surface perpendicular to the magnetic flux density, as shown in **Figure 5.98**. The magnetic flux density B produced by the measured current I produces a Hall voltage according to **Equation (5.44)**:

$$V_{\text{out}} = K_H \frac{I_H B}{d} = \left(K_H \frac{I_H}{d} \frac{\mu_0}{2\pi a} \right) I \quad [\text{V}]. \tag{5.79}$$

In this relation, I_H is the bias current through the Hall element (see **Figure 5.38**), d is the thickness of the Hall element, and μ_0 is the permeability of the material in which the Hall element is embedded (assuming that it is nonmagnetic). K_H is the Hall coefficient.

EXAMPLE 5.20 A current sensor

Any current in a long conductor produces a magnetic flux density as described in **Equation (5.17)**. This magnetic flux density can be measured using a Hall element and by doing so produce a current sensor that indicates the current without the need to cut the conductor and measure the current in the conventional manner. The current sensor is shown in **Figure 5.98**. The only requirements are that the Hall element and the conductor are in a plane so that the magnetic flux density of the conductor, which is circumferential to the conductor, is perpendicular to the surface of the Hall element (or that the sensor be properly calibrated if that condition cannot be satisfied). The Hall element is embedded in a nonconducting ring that fits snuggly over the conductor and holds the Hall element in the position shown. (The ring is not shown, as it only has a mechanical function and does not affect the Hall element reading.) Using a small Hall element with a Hall coefficient of 0.01 m^3/A/s at a fixed distance of 10 mm from the center of the conductor, calculate the response of the sensor for currents between 0 and 100 A using the Hall element and biasing in **Example 5.7**.

Solution: First we calculate the range of the magnetic flux density for the given current using **Equation (5.17)**:

$$B = \frac{\mu_0 I}{2\pi r} = \frac{4\pi \times 10^{-7} I}{2\pi \times 0.01} = 2 \times 10^{-5} I \quad [\text{T}].$$

This is then introduced into **Equation (5.44)**:

$$V_{\text{out}} = K_H \frac{I_H B}{d} = K_H \frac{I_H \times 2 \times 10^{-5} I}{d} = 0.01 \times \frac{5 \times 10^{-3}}{0.1 \times 10^{-3}} 2 \times 10^{-5} I = 10^{-5} I \quad [\text{V}].$$

The same result may be obtained directly from **Equation (5.79)**. For the maximum conductor current of 100 A, this produces an output of 1 mV. Therefore the output will vary linearly from 0 to 1 mV for a current varying between 0 and 100 A. The sensitivity is 10^{-5} V/A, clearly requiring amplification for practical use.

5.11 | PROBLEMS

Capacitive sensors and actuators

5.1 **Capacitive position sensor**. A position sensor is made as follows: Two tubes with very thin walls are placed one inside the other so that they are concentric, separated by a Teflon layer, and can move in or out between two preset limits x_1 and x_2, as shown in **Figure 5.99**. The tubes are $L = 60$ mm long.

a. Calculate the minimum and maximum capacitance between the two tubes, assuming that an electric field may only exist in that area in which the tubes overlap.

b. Show how this device can be used as a sensor; that is, calculate its sensitivity for $a = 4$ mm, $b = 4.5$ mm, $x_1 = 10$ mm, $x_2 = 50$ mm, and the relative permittivity of Teflon $\varepsilon_r = 2.25$.

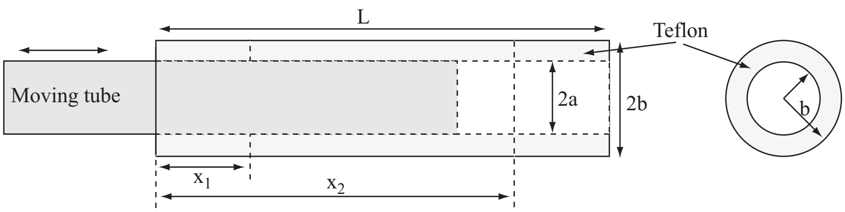


FIGURE 5.99 ■ Capacitive position sensor.

5.2 **Capacitive temperature sensor for water.** The temperature of water can be measured directly using a capacitive sensor based on the fact that the permittivity of water is highly dependent on temperature. As temperature changes from 0°C to 100°C, the relative permittivity of water, ε_w, changes from 90 to 55.

a. Given a sensor as shown in **Figure 5.100**, calculate its sensitivity assuming the permittivity change is linear with temperature.

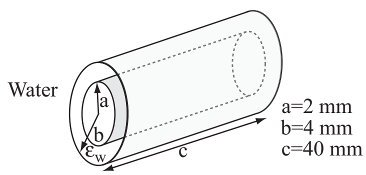


FIGURE 5.100 ■ Capacitive temperature sensor for water.

b. With a digital capacitance meter capable of a resolution of 0.2 pF, what is the resolution of the sensing system in terms of temperature?

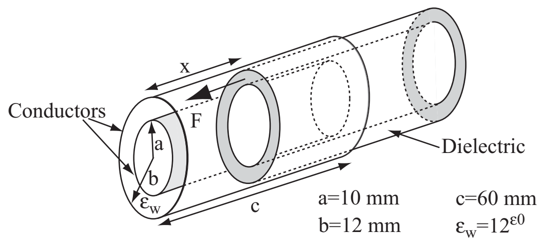
5.3 **A low-force capacitive actuator.** A simple capacitive actuator can be made using the configuration in **Figure 5.101**. Two coaxial tubes form a capacitor. The space between them contains a hollow tube made of Teflon that is free to move between the two cylinders as shown. Assume the moving part is at an arbitrary location x from the edge of the capacitor and the dimensions are as shown in the figure:

a. Calculate the force the actuator can exert as a function of the applied voltage connected across the cylinders (positive to the outer cylinder, negative to the inner cylinder).

b. Show from physical considerations that the motion can only be inward, regardless of polarity of the voltage applied.

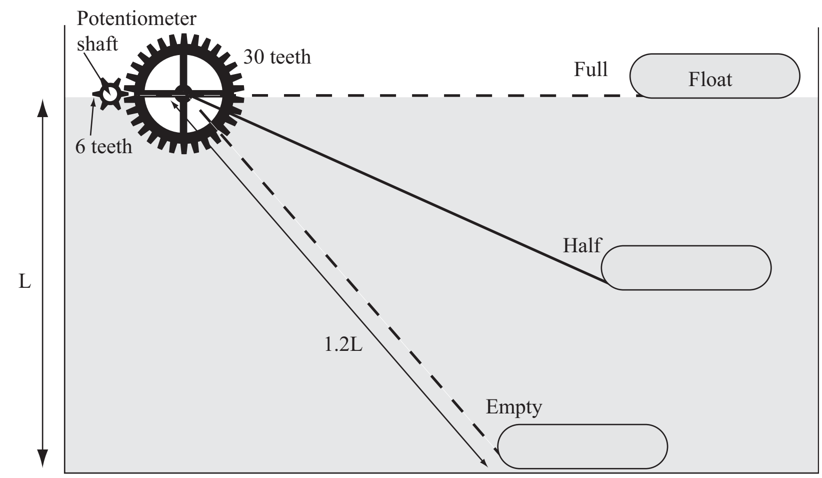
Note: The capacitance of coaxial capacitors can be calculated directly but, as a first approximation, they may be treated as parallel plate capacitors using **Equation (5.2)**, with S taken as the average between the areas of the outer and inner conductors, especially if the inner radius is not too small. The same applies to the electric field in the capacitor.

FIGURE 5.101 ■ Capacitive force actuator.



5.4 **Resistive (potentiometer) fuel tank gauge.** A fuel tank gauge is made as shown in **Figure 5.102**. The rotary potentiometer is linear with a resistance $R = 100\ \mathrm{k\Omega}$ distributed over a 330° strip (i.e., the potentiometer can rotate 330°). The float is connected to the shaft of a gear with 30 teeth and that in turn rotates the potentiometer

FIGURE 5.102 ■ Resistive fuel gauge.



through a gear with 6 teeth. The mechanical linkage is set so that when the tank is full, the potentiometer resistance is zero. The float is shown in three positions to show how the resistance changes (increases from zero at full tank).

a. Calculate the resistance reading for empty, ¼, ½, ¾ and full tank.

b. Using the points in (a), find a linear best fit and calculate the maximum non-linearity of the sensor. Why is the curve nonlinear?

5.5 **Corrosion rate sensor.** Structures that are subject to corrosion require a means of determining the corrosion rate so that action can be taken before the structure becomes dangerous. One way of doing so is to expose a thin wire of the same material as the structure to identical conditions by placing it in the same location as the monitored structure. The corrosion rate is often defined as corrosion depth in millimeters per year (mm/year). A corrosion rate sensor is made of a steel wire L m long and d mm in diameter. Its resistance is monitored continuously and correlated directly with the corrosion rate.

a. Find a relation between the rate of corrosion in millimeters and resistance measured. Assume corrosion is uniform around the circumference and that the corrosion products do not contribute to resistance.

b. Discuss how variations in temperature may be accounted for.

5.6 **Tide level sensor.** In an attempt to monitor maximum and minimum tide levels, a simple sensor is made as follows (**Figure 5.103**). A metal tube of inner diameter $a = 100$ mm is coated on the outside with an isolating layer of paint so that seawater cannot come in contact with the outer surface of the tube. A metal cylinder of outer diameter $b = 40$ mm is placed inside the larger tube and, using a few insulating spacers, held so that the two tubes are coaxial. Now the assembly is sunk into the sea. A series of small holes at the bottom allow seawater into the space between the tubes. A 1.5 V battery is connected through an ammeter as shown.

a. If the maximum and minimum water levels are as shown, and the conductivity of seawater is $\sigma = 4$ S/m, calculate the maximum and minimum (range) reading of the ammeter. Assume conductors are perfect conductors and both the air and sea bottom are insulators.

b. Calculate the sensitivity of the sensor.

c. Calculate the resolution for a digital ammeter that measures in steps of 1 mA.

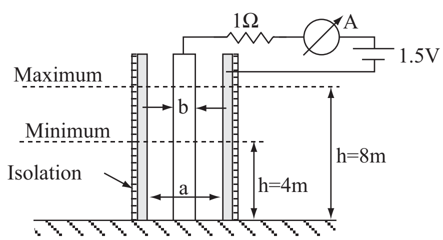


FIGURE 5.103 ■ Tide level sensor.

5.7 **Resistive position sensor.** A position sensor is made as shown in **Figure 5.104**, where the inner plate can slide back and forth over a distance of 10 cm. Both the stationary and moving sections are made of graphite and are in intimate contact. The length of the moving bar is the same as the length of the stationary section (15 cm). The conductivity of graphite is $\sigma = 10^2$ S/m. Given the dimensions in the figure,

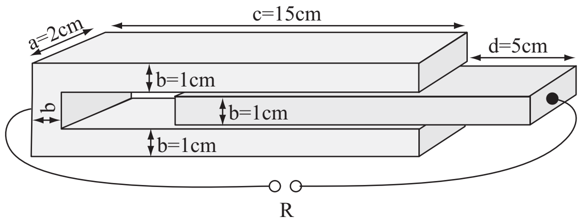


FIGURE 5.104 ■ A simple resistive position sensor.

a. Calculate the relation between position d and resistance R measured between the two.
b. Calculate the maximum and minimum resistance possible (i.e., for $d = 1$ cm and for $d = 11$ cm).
c. Calculate the sensitivity of the sensor.

Magnetic sensors and actuators

5.8 **Induction proximity sensor (induced emf).** In some applications it is important to sense displacement or proximity without incurring friction and losses. One possible solution is shown in **Figure 5.105**. Two solenoids of length L each are placed inside one another as shown. The radii of the solenoids are a and b and each has N turns uniformly wound over the length L. A sinusoidal current with amplitude I_0 and frequency f flows in the outer solenoid, which is stationary. $I_0 = 0.1$ A, $f = 1000$ Hz, $L = 10$ cm, $a = 10$ mm, $b = 12$ mm, and $N = 400$ turns. The coils are shown in axial cross section. The ⊗ in the turns indicates a current flowing in, whereas a ⊙ indicates a current flowing out of the plane of the cut.

a. What is the emf induced in the inner, movable solenoid when the two coils overlap a distance x? Assume that $L >> b$, which allows you to say that the field is approximately constant within the solenoid and zero outside and neglect the effects of the ends of the outer solenoid. The moving coil core is nonmagnetic (permeability equals that of space).
b. What is the sensitivity of the sensor (voltage RMS per unit length) if the core is replaced with iron with relative permeability μ_r?

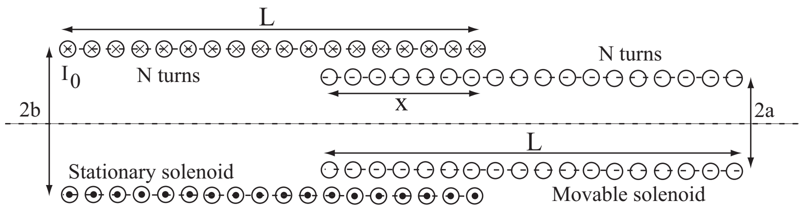


FIGURE 5.105 ■ A magnetic position sensor.

5.9 **Invisible fence for dogs.** In this type of system, a wire is buried under the surface and a current at a given frequency passes through the wire. The dog wears a small unit made of a pickup coil and electronics that deliver a high-voltage pulse to the dog through a couple of electrodes pressed against its skin.

The pulse is not harmful, but it is painful enough to cause the dog to keep away. In an invisible fence, the wire carries a 0.5 A (amplitude) sinusoidal current at 10 kHz. The dog carries a sensor made as a coil with 1500 turns and 30 mm in diameter.

a. If the detection level is set at 200 μV RMS (i.e., the level at which the dog will receive a correction pulse), what is the furthest distance from the wire the dog will "feel" the presence of the fence?
b. What are the conditions necessary to obtain the result in (a)?

Hall effect sensors

5.10 **Hall effect in conductors.** The Hall effect in conductors is rather small and can be calculated from **Equation (5.43)**. To see what the order of magnitudes of the Hall coefficient and Hall voltage are, consider a Hall sensor made by deposition of gold on a substrate. The sensor itself is 2 mm × 4 mm and is 0.1 mm thick. The free charge density of gold is 5.9×10^{28} electrons/m^3. Calculate the Hall coefficient and the sensitivity of the sensor in sensing magnetic flux densities. Assume the magnetic field is perpendicular to the plate and a current of 15 mA flows along the long dimension of the plate.

5.11 **Hall effect in silicon.** A Hall element is made in the form of a small silicon wafer 1 mm × 1 mm and 0.2 mm thick. An *n*-type silicon is used with a majority carrier density of 1.5×10^{15} carriers/cm^3, whereas the intrinsic carrier density is 1.5×10^{10} carriers/cm^3. The mobility of holes is 450 cm^2/V·s and that of electrons is 1350 cm^2/V·s. Calculate

a. The Hall coefficient of the Hall element.
b. The sensitivity of the Hall sensor in sensing magnetic fields given a fixed current of 10 mA across the sensor. Assume the magnetic flux density is applied perpendicular to the silicon plate.
c. Suppose that the intrinsic material is used to make the Hall element. Use the dimensions of the wafer to calculate the resistance of the Hall element and explain why this Hall element is not a practical device.

5.12 **Zero Hall coefficient semiconductor.** If a semiconducting device must operate in high magnetic fields, and if the purpose is not to measure the magnetic field, the Hall voltage may be detrimental to the operation of the device. Under these conditions it may be useful to dope the semiconductor so as to produce a zero Hall coefficient. Given the mobility of holes in silicon as 450 cm^2/V·s and that of electrons as 1350 cm^2/V·s, what is the ratio of *n* to *p* densities required?

5.13 **Power sensor.** A power sensor is built as shown in **Figure 5.40** as part of a DC power meter for a small electric vehicle. A small ferromagnetic core with very large permeability is used and the Hall element is placed in the small gap shown. The magnetic core saturates at a magnetic flux density of 1.4 T. The Hall element has a Hall coefficient of $K_H = 0.018$ m^3/A/s and is $d = 0.4$ mm thick. The gap is slightly larger, at $l_g = 0.5$ mm, so that the Hall element fits snuggly in the gap. Neglect the resistance of the Hall element and of the coil.

a. Assuming N turns in the coil, a resistance R, and a resistive load R_L, find the expression relating the Hall voltage and the power in the load (this is the transfer function of the sensor).

b. Find the maximum power the sensor can sense, assuming the line voltage is constant at 12 V and the coil contains $N = 100$ turns.

c. What is the sensor's reading under the conditions in (b) if the Hall element operates at a current of 5 mA?

d. What is the sensitivity of the power sensor?

5.14 **Carrier density sensor.** The Hall effect can be used in many ways, and not all of them involve sensing of magnetic fields. Consider the measurement of carrier density in a metal alloy. To perform the measurement, the metal is cut into a rectangular plate 25 mm long, 10 m wide, and 1 mm thick. The metal is placed between two neodymium-iron-boron (NeFeB) magnets, producing a constant 1 T magnetic flux density through the disk. A current of 1 A passes through the plate and the voltage across the plate is measured as 7.45 μV (**Figure 5.106**).

a. Show the polarity of the voltage given the direction of the magnetic flux density and that of the current.

b. Calculate the carrier density in the metal alloy.

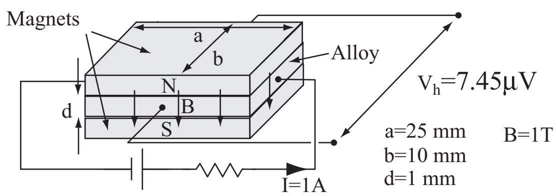


FIGURE 5.106 ■ Carrier density sensor.

5.15 **Doping concentration in *n*-type silicon**. It is required to estimate the carrier density in an *n*-type silicon sample as a means of production control. A sample 2 mm long, 1 mm wide, and 0.1 mm thick is prepared and connected as in **Figure 5.38** or **Figure 5.106**, carrying a current of 10 mA. The sample is placed between the poles of a neodymium-iron-boron (NeFeB) permanent magnet producing a magnetic flux density of 0.8 T perpendicular to the surface of the sample and a 80 μV Hall voltage is measured across the sample. Calculate the carrier concentration in the sample assuming the majority carriers dominate. Electrons in silicon have a mobility of 1350 $cm^2/V{\cdot}s$.

Magnetohydrodynamic sensors and actuators

5.16 **Magnetohydrodynamic actuator for submarine propulsion.** A submarine is designed to move using a magnetohydrodynamic pump. The pump is made in the form of a channel that is $b = 8$ m long, $d = 0.5$ m high, and $a = 1$ m wide (**Figure 5.48**). The magnetic field is constant everywhere throughout the channel in the direction shown and equal to $B = 1$ T. A current $I = 100$ kA passes through the electrodes shown and through the water between the electrodes, generating a force on the water as shown. The conductivity of seawater is 4 S/m.

a. Calculate the force produced by the pump.

b. Calculate the potential required between the two electrodes.

5.17 **Magnetohydrodynamic generator.** A magnetohydrodynamic generator is made in the form of a channel 1 m long with a cross section of 10 cm × 20 cm (**Figure 5.107**). The magnetic flux density produced between the poles on the narrow side is 0.8 T. A jet combustor is used to drive combustion gases through the channel and the gases are seeded with conducting ions to produce an effective conductivity of 100 S/m in the channel. The combustor drives the gases at 200 m/s, effectively making the gas that moves through the channel a plasma.

a. Calculate the output voltage of the generator (*emf*).

b. Calculate the maximum power it can generate if the output voltage cannot change by more than 5%.

c. Indicate qualitatively how this method might be used for sensing and what the quantities are it can sense.

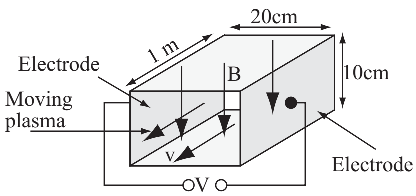


FIGURE 5.107 Magneto-hydrodynamic generator.

5.18 **Magnetic flowmeter.** A magnetic flowmeter can be built as follows (**Figure 5.108**). The fluid flows in a square cross-section channel. On two opposite sides there are two coils that produce a constant magnetic flux density B_0 pointing down. The fluid contains positive and negative ions (Na^+, Cl^-, etc.). The magnetic flux density is assumed to be constant and uniform in the channel. This forces the positive and negative charges toward the two opposing electrodes.

a. Calculate the potential difference for a flowing fluid as a function of its velocity v and show its polarity. The cross-sectional area of the channel is $a \times a$ [m^2].

b. Calculate the sensitivity of the sensor to flow. Flow is measured as volume per second (i.e., m^3/s). Discuss realistic ways to improve the sensitivity of the device.

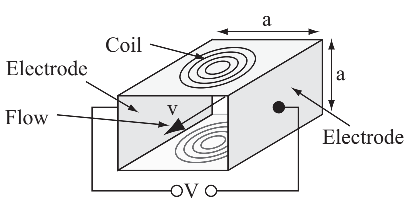


FIGURE 5.108 Magnetic flowmeter.

5.19 **Magnetic gun (magnetic force).** A magnetic gun is made as shown in **Figure 5.109**. Two cylindrical conductors of diameter a (rails) and separated a distance d between centers are shorted by a conducting projectile that is free to move along the rails. When a current is applied, the magnetic flux density produced by the current in the conductors generates a force on the projectile.

a. Given $a = 10$ mm, $d = 40$ mm, and $I = 100{,}000$ A, what is the force on the projectile?

b. If the projectile has a mass of 100 g, calculate the acceleration and the exit velocity of the projectile for rails 5 m long with the projectile starting at one end and exiting the other.

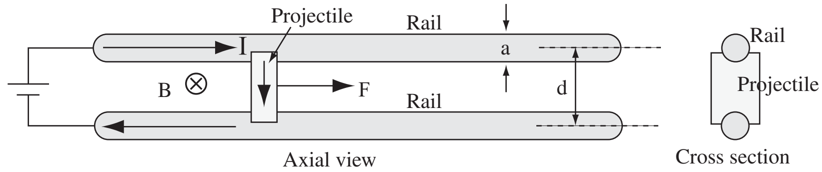


FIGURE 5.109 ■ The rail gun. (a) Axial view. (b) Cross section through the projectile.

5.20 **Magnetic density sensor.** The density of a fluid can be sensed using a magnetic sensor as follows. A sealed float (i.e., a closed container) is equipped with an amount of iron on its bottom and a coil at the bottom of the container containing the fluid is driven with a current I, as shown in **Figure 5.110**. The current in the coil is increased until the float is suspended in the fluid with its top at the surface of the fluid. The density of the float, including the iron, is known as ρ_0 and its volume is V_0. The force the coil exerts on the float equals kI^2, where k is a given constant that depends on the amount of iron in the float, the size of the coil, and distance to the float. The current in the coil is the measured quantity.

a. Find the transfer function of the sensor, that is, find the relation between the density of the fluid, ρ, and the measured current, I.

b. Calculate the sensitivity of the sensor.

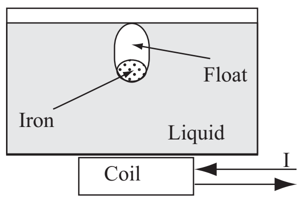


FIGURE 5.110 ■ Magnetic density sensor.

Magnetostrictive sensors and actuators

5.21 **Optical fiber magnetometer.** An optical fiber magnetometer is made of two fibers, each 100 m long. One fiber is coated with nickel and an infrared light emitting diode (LED) emitting at 850 nm is used as the source for both fibers. Light propagates along each fiber with a phase constant $\beta = 2\pi/\lambda$ [rad/m], where λ is the wavelength of the light in the fiber. That is, for every 1 m length of the fiber, the phase changes by $2\pi/\lambda$ radians. At the end of the fibers, the phases of the two signals are compared. The signal propagating through the coated fiber will have a lower phase since the sensed magnetic flux density causes the nickel coating, and hence the fiber, to contract (nickel is a magnetostrictive material; see **Table 5.5**). The device is shown in **Figure 5.111**. Assuming that the phase detector can detect a phase difference of 5°, calculate the lowest magnetic flux density detectable using this magnetometer.

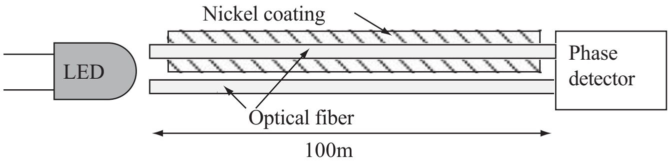


FIGURE 5.111 ■ Optical fiber magnetometer.

Voice coil actuators

5.22 **Voice coil actuator.** A cylindrical voice coil actuator is shown in **Figure 5.112**. A permanent magnet, magnetized radially, is placed on the outer surface of the gap with a free moving coil on the inner cylinder that can move back and forth. The mass of the coil is $m = 50$ g, the current in the coil has amplitude $I = 0.4$ A and frequency $f = 50$ Hz, the number of turns in the coil is $N = 240$, and the magnetic flux density due to the permanent magnet is $B = 0.8$ T. The coil has an average diameter of 22.5 mm. This device is used as a positioner or as the drive in a vibrating pump. The activating mechanism is connected directly to the coil.

a. Calculate the force on the coil.

b. If the cylinder has a restoring constant $k = 250$ N/m (provided by a spring), calculate the maximum displacement of the moving piece.

c. Calculate the maximum acceleration of the coil/cone assembly.

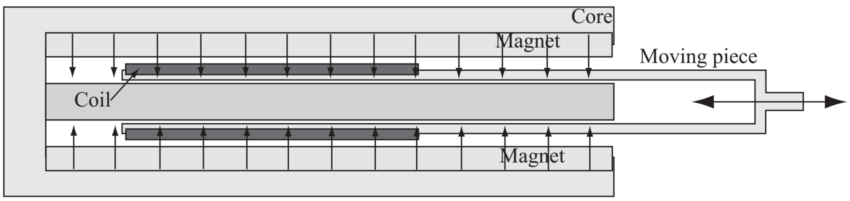


FIGURE 5.112 ■ A voice coil actuator. The magnetic field is indicated by the vertical arrows.

Motors as actuators

5.23 **A simple DC motor.** A simplified form of a motor is shown in **Figure 5.113**. The rotor and stator are separated by a gap of 0.5 mm. At a given time, the coil, which is embedded into the rotor in a groove, made of 100 turns, and carrying a current $I = 0.2$ A, is oriented as shown in the figure. The current is into the page on the top of the rotor and out of the page on the bottom of the rotor. The rotor is a cylinder of radius $a = 2$ cm and depth $b = 4$ cm. The radius of the coil in the rotor may be assumed to be the same as that of the rotor.

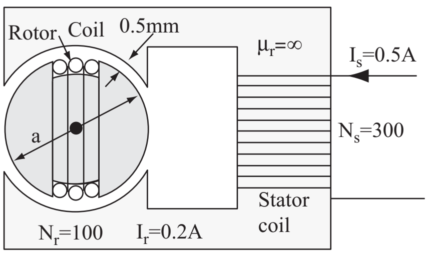


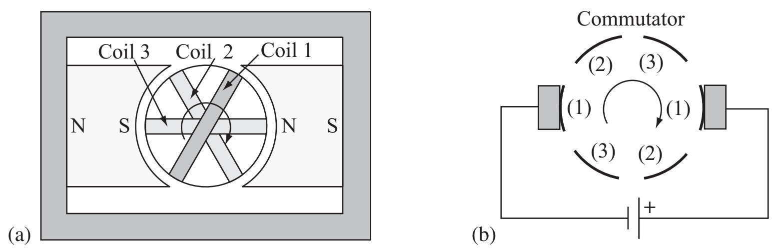
FIGURE 5.113 ■ A simplified DC motor.

a. Show the direction of rotation of the rotor for the configuration in **Figure 5.113**.

b. For the given currents, calculate the maximum torque this motor is capable of.

5.24 Permanent magnet DC motor with three coils. A permanent magnet DC motor has three coils in the rotor with a six-strip commutator, as shown in **Figure 5.114a**. Each of the three coils is permanently connected to a pair of strips on the commutator, as indicated. Assume the magnetic flux density in the gap is uniform and each coil is, in turn, horizontal when the brushes are at the center of the corresponding commutator strips (**Figure 5.114b** shows coil 1 in that position). Each pair of strips occupies one-third of the circle shown in **Figure 5.114b**. The radius of the rotor is 30 mm, the length 80 mm, and the permanent magnets produce a magnetic flux density of 0.75 T in the gap and throughout the rotor. The current in the coils is 0.5 A and each coil is made of 120 turns. Calculate and plot the torque as a function of time for a full cycle of the rotor.

FIGURE 5.114 ■ Permanent magnet DC motor with three coils.

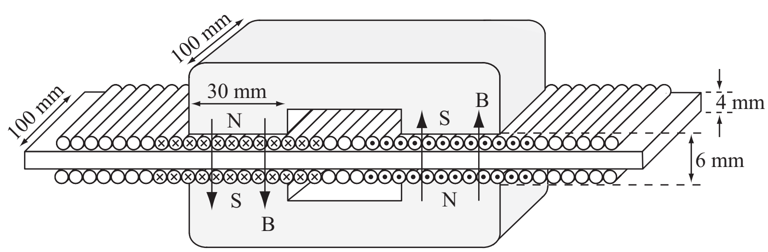


5.25 Linear DC motor. A linear DC motor is shown in **Figure 5.115**. A permanent magnet generates a constant magnetic flux density of 0.5 T in the gap between its poles. An iron plate 4 mm thick is placed between the poles of the magnets, leaving enough freedom for the magnets to move, and a coil with 1000 turns/m is wound on the iron plate. The gap on each side of the iron plate is 1 mm and the current-carrying wires shown are bonded to the plate. Wires with an ⊗ carry a current into the page, those with a ⊙ carry a current out of the page, and those without a symbol do no carry a current. As the magnet assembly moves under the influence of the magnetic forces, the currents in the conductors are shifted so that at all times the relation between currents and magnets remains the same. The plate with the conductors is stationary.

a. Calculate the force on the magnet for a current of 5 A and show the direction of motion of the magnet assembly for the current shown in the figure.

b. Show how this configuration may be changed into a stepping motor without changing the geometry and define its possible step sizes based on your configuration.

FIGURE 5.115 ■ Permanent magnet linear motor.



Stepper motors

5.26 General relations in stepper motors. The number of teeth in the rotor and stator of a stepper motor define the number of steps the motor is capable of. Show in general that if n is the number of teeth in the stator and p is the number of teeth in the rotor, then the following applies:

a. The larger n is for a fixed value p, the larger the step size.
b. The smaller the difference $n-p$, the smaller the step size.
c. The larger the numbers n and p, the smaller the step size.
d. Discuss the limitations on n and p.

5.27 High angular resolution stepper motor. A variable reluctance stepper motor is proposed with six poles in the stator and 6 teeth in each pole and 50 teeth in the rotor (usually the number of teeth in the rotor is less than in the stator because its diameter is smaller, but this is not a requirement, and in variable reluctance motors the opposite is just as practical).

a. Calculate the angle per step and number of steps per revolution.
b. Explain why a motor with the number of teeth indicated here is not likely to be produced even though it is entirely possible to do so.

5.28 Noninteger number of steps per revolution. Normally a stepper motor is designed with an integer number of steps per revolution. However, one can envision cases of noninteger number of steps, such as in cases where a specific step size is needed (say, 4.6°). Consider a motor with eight poles in the stator with 6 teeth per pole and 38 teeth in the rotor. Calculate the step size and the number of steps per revolution.

5.29 Linear stepping motor. A linear stepping motor similar to the one in **Figure 5.83** has N teeth/cm in the stator and M teeth/cm in the slider.

a. Derive a relation that allows calculation of the step size in millimeters.
b. What is the step size for a linear motor with 10 teeth/cm in the stator and 8 teeth/cm in the slider?

5.30 Force in a magnetic valve. Calculate the force on the movable part in the configuration shown in **Figure 5.116**. Edge effects in the air gaps are neglected and the flux density is assumed to be constant and perpendicular to the surfaces of the gaps. The permeability of the core may be assumed to be very large, number of turns $N = 300$, $I = 1$ A, $a = 25$ mm, $b = 10$ mm, $l = 2$ mm, and $L = 20$ mm.

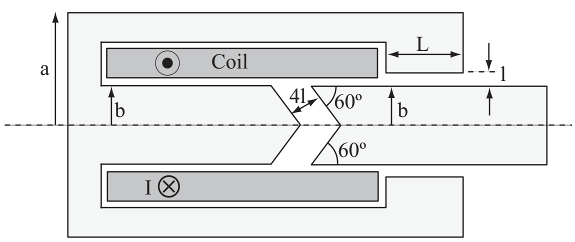


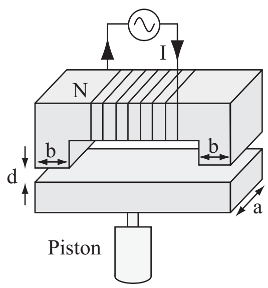
FIGURE 5.116 ■ Structure of a magnetic valve.

5.31 Solenoid actuator in a paint sprayer. A pump used to spray paint in an airless sprayer is made as shown in **Figure 5.117**. The principle is as follows: when a current is applied to the coil, the gap closes, and it opens when the current is switched off. If an alternating current is applied, the armature vibrates since the field is sinusoidal and passes through zero twice each cycle. This moves a piston back and forth a short distance, sufficient to activate a piston to pump the paint from its reservoir and spray it through an orifice. **Figure 5.117** shows a simplified form of the structure without the spray mechanism itself. In an actual device, the armature is hinged on one side and has a restoring spring. The coil contains $N = 5000$ turns and carries a sinusoidal current at 60 Hz with amplitude $I = 0.1$ A. The permeability of iron is assumed to be very large and the gaps are $d = 3$ mm. The dimensions $a = 40$ mm and $b = 20$ mm define the surface area of the poles that form the gap. Assume all flux is contained in the gaps (no leakage of flux outside the area of the poles).

a. Calculate the force exerted by the moving piece on the piston.

b. What is the force if the gap is reduced to 1 mm?

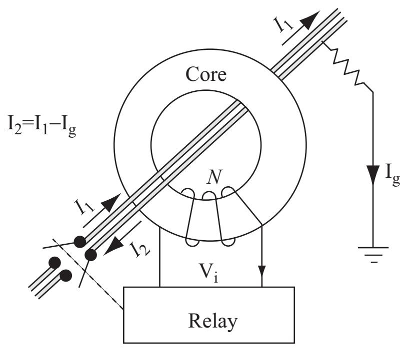
FIGURE 5.117 ■ Actuator for an airless sprayer.



Voltage and current sensors

5.32 Ground fault circuit interrupt (GFCI). An important safety device is the GFCI (also called a residual current device (RCD)). It is intended to disconnect electrical power if current flows outside of the intended circuit, usually to ground, such as in the case when a person is electrocuted. The schematic in **Figure 5.118** shows the concept. The two conductors supplying power to an electrical socket or an appliance pass through the center of a toroidal coil or a Rogowski coil.

FIGURE 5.118 ■ Principle of a GFCI sensor.



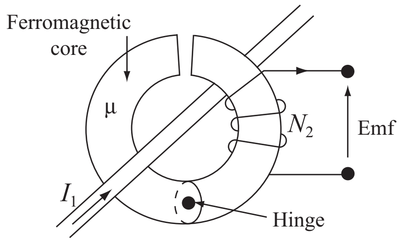


FIGURE 5.119 ■ Principle of the clamping ammeter.

Normally the currents in the two conductors are the same and the net induced voltage due to the two conductors cancel each other, producing a net zero output in the current sensor. If there is a fault and current flows to ground, say a current I_g, the return wire will carry a smaller current and the current sensor produces an output proportional to the ground current I_g. If that current exceeds a set value (typically 30 mA), the voltage induced causes the circuit to disconnect. These devices are common in many locations and are required by code in any location in close proximity to water (bathrooms, kitchens, etc.). Consider the GFCI shown schematically in **Figure 5.118**. The device is designed to operate in a 50 Hz installation and trip when the output voltage is 100 μV RMS. For a toroidal coil with average diameter $a = 30$ mm and a cross-sectional diameter of $b = 10$ mm,

a. Calculate the number of turns needed if a Rogowski coil is used and the device must trip at a ground current of 30 mA.
b. Calculate the number of turns needed to trip at a current of 30 mA if a ferromagnetic torus with relative permeability of 2000 is used as a core for the coil.

5.33 **Clamping ammeter.** When measuring the current in high current conductors without the need to cut the conductor and insert a regular ammeter, one uses a toroidal coil with the wire in which the current is measured passing through the toroid. To accomplish this, the torus is hinged so it can be opened and closed around the wire (see **Figure 5.119**). The *emf* in the coil on the torus is measured and related to the current in the wire.

a. Find the relation between current (RMS) and emf measured (RMS).
b. What is the maximum *emf* (peak value) for a sinusoidal current of amplitude 10 A at 60 Hz for a torus with inner radius $a = 2$ cm, outer radius $b = 4$ cm, thickness $c = 2$ cm, $N = 200$ turns, and relative permeability $\mu_r = 600$. The torus has rectangular cross section.

5.34 **AC sensor.** One type of commercially available current sensor is made as a simple closed magnetic core, typically rectangular, as shown in **Figure 5.120**.

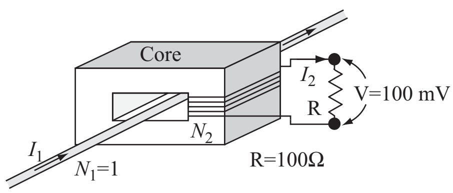


FIGURE 5.120 ■ Current sensor.

The current in a conductor is sensed by threading the wire through the central opening (this type of sensor is typically used in fixed installations) and the induced voltage in a coil wound uniformly on the core, or the current through it is used to measure the current in the wire. For a full scale of 100 A (RMS), the sensor shown is designed so that the induced voltage is 100 mV (RMS) across a load of 100 Ω. Calculate the number of turns required in the secondary coil assuming the current sensor operates as an ideal transformer.